

DIANA TEKNOLOGIA presenta **XerKa**, una innovador software para la localización, clasificación y suministro de la información requerida por el usuario de forma efectiva.

Las herramientas de búsqueda y análisis de información incorporadas hacen posible el tratamiento en varios idiomas de millones de documentos para una correcta interpretación y clasificación de los mismos.

El novedoso sistema de recuperación realiza una localización rápida y efectiva de la información previamente clasificada, lo que convierte a **XerKa** en la solución ideal como

sistema de vigilancia
tecnológica y de
mercado

clasificador inteligente
de contenidos temáticos

Una Solución eficaz y avanzada para la búsqueda, categorización y recuperación de Contenidos Temáticos

en redes digitales
extensas de información

XerKa es un **Sistema de Información** que se encarga de Proporcionar al Usuario un **Acceso Directo** a la Información que necesita. **Explora ininterrumpidamente** cualquier red digital de información accesible (Internet, redes privadas, etc.) seleccionando los sitios que puedan ser relevantes para los usuarios del sistema.

Filtra, Categoriza y Clasifica automáticamente las páginas y documentos encontrados dentro de un espacio temático Multidimensional accesible tanto directamente como a través de un Sistema de Alertas.

Tecnología software a servicio de la Vigilancia Tecnológica / Inteligencia Competitiva

Amplia precisión y cobertura en la búsqueda de información

- Búsqueda enfocada (a sitios de interés)
- Argumentada (por el espacio temático)
- Potente (múltiples robots)
- Selectiva (filtra lo interesante)

Exploración exhaustiva del contenido de los sitios enfocados (pln)

- Análisis completo del contenido del documento
- Filtrado por tipo de documento
- Análisis multilingüe de la información

Indexación multidimensional de los documentos clasificados

- Reindexación dinámica de la base documental

Múltiples posibilidades para la explotación de la información clasificada

- Rosa de los temas*: consulta a la base y acceso directo al documento
- XerKa Alert*: Servicio de alertas
- Index*: Búsqueda complementaria por lemas y palabras
- Xerka-Cluster*: Agrupación de documentos por similitud en su contenido.

Tecnología punta

Procesamiento del lenguaje natural.

Tratamiento multilingüe de la información: español, euskera, inglés y francés

XerKa analiza el contenido textual de los documentos y presenta (en su idioma original) los documentos seleccionados como respuesta a cada consulta del usuario.

Trabajar de forma coordinada en cuatro idiomas permite a **XerKa** superar la problemática con la que se encuentra el usuario al utilizar buscadores por términos: el desconocimiento de la palabra clave adecuada y su traducción a otras lenguas.

Indexación multidimensional

XerKa indexa los documentos en función del contenido de los mismos de forma que queda **clasificado por múltiples criterios**.

Esto permite al usuario realizar las consultas en busca de documentos que traten varios temas a la vez y obtener una **respuesta rápida y efectiva**.

Aplicación práctica de XerKa como sistema de vigilancia en la empresa

En la empresa, una vez interiorizado el valor de la vigilancia y su necesidad, se plantea el paso de la teoría a la práctica. ¿Se puede sistematizar la explotación de información? ¿Se puede automatizar? La respuesta es sí. **El objetivo de XerKa**, como herramienta de vigilancia, **es automatizar las diferentes fases del proceso de VT/IC y proporcionar al usuario la información precisa para su explotación en los procesos de inteligencia competitiva**.

El siguiente cuadro resume las funcionalidades de **XerKa** vinculadas a las fases de búsqueda, análisis y difusión de la información relevante en un proceso de VT/IC

Fases del proceso VT-IC	Puntos clave de Xerka
Identificar y analizar las necesidades	Definición de Espacios Temáticos por usuario o grupos de usuarios: trabajo base de consultoría en el que se estudian las necesidades de cada empresa y se generan los ejes temáticos de Xerka
Obtener la información	Captación automática mediante robots que realizan una búsqueda inteligente de información. Xerka realiza una exploración multilingüe in-interrumpida en Internet , logrando una gran rapidez, cobertura y precisión en la obtención de resultados.
Analizar, procesar y valorizar la información	Analiza la información por su contenido (procesamiento del lenguaje natural) y la clasifica por múltiples criterios (indexación multidimensional) Revisión continua y automática de los cambios en los documentos y actualización continua de Nuevos Términos para describir los espacios temáticos de interés-
Difundir los resultados	Proporciona un acceso directo a la información a través de la interfaz <i>Rosa de los Temas</i> donde el usuario realiza sus consultas . Además, dispone de un módulo opcional para la generación de alertas (XerKa-Alert) , la localización de documentos por concurrencia de palabras clave (<i>Index</i>) y la agrupación de documentos por similitud de contenido (<i>XerKa-Cluster</i>)

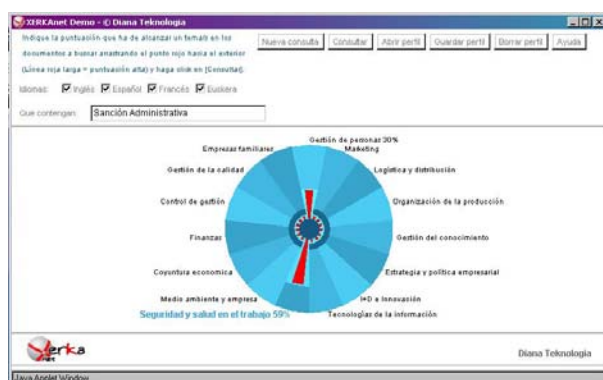
¿Cómo explotar la base documental alimentada por XerKa?

Rosa de los Temas



En la interfaz de consulta básica ("Rosa de los Temas") el usuario solicita información relativa a uno o varios ejes temáticos y el sistema responde con un listado de documentos.

La longitud cada eje representa la importancia (relevancia) que debe tener ese tema en los documentos. El usuario simplemente tiene que alargar o acortar cada eje para realizar la consulta.



Por ejemplo: Documentación sobre seguridad y salud en el trabajo que contenga información sobre organización de equipos de trabajo.

- Seguridad y salud en el trabajo 59ptos
- Gestión de personas 30ptos

La posibilidad de combinación de varios ejes es lo que verdaderamente enriquece al sistema, ya que permite contextualizar la información a un determinado ámbito.

El sistema mostrará una **lista de documentos**, que han recibido la puntuación solicitada por el usuario. La lista incluye: un extracto de cada documento, una relación de términos descriptores del mismo y un enlace al sitio donde está alojado.

Basta con hacer click en el documento seleccionado para **acceder a él en su formato original**.

Además el usuario puede acceder los módulos accesorios de **XerKa** para:

a) Complementar la consulta con una búsqueda por lemas.

XerKa Index

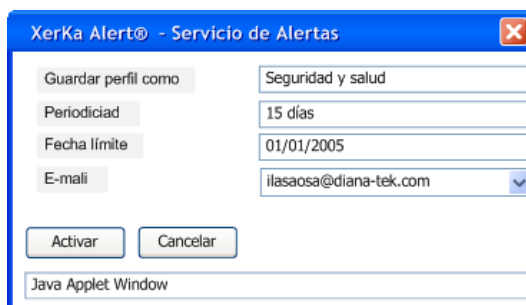


Por ejemplo: "sanción administrativa".

El sistema mostrará todos los documentos que traten sobre seguridad y salud en el trabajo y organización de equipos en los que aparezcan "sanción administrativa" (cualquier forma superficial del término).

b) Generar y distribuir alertas con los nuevos documentos incorporados a la BBDD.

XerKa Alert

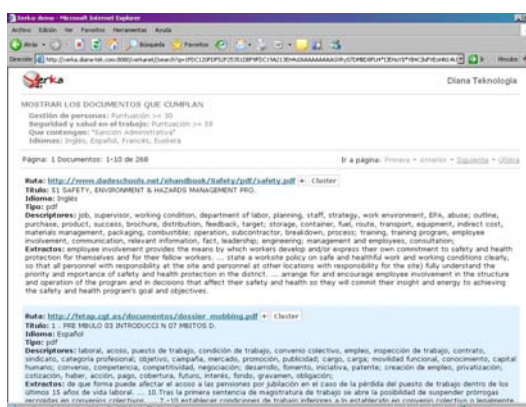


XerKa-Alert es un sistema programado de difusión selectiva de información (DSI).

Realiza de forma periódica la consulta de un perfil creado por el usuario y le envía un correo con el extracto de los documentos nuevos incorporados que tienen el contenido requerido.

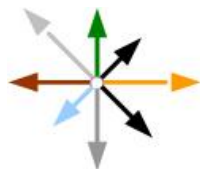
c) Agrupar documentos cuyo contenido es similar al seleccionado.

XerKa Cluster



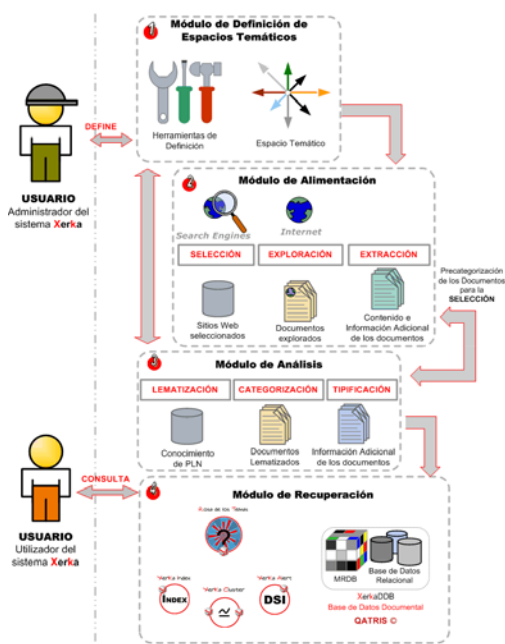
Aproximación técnica: módulos que integran XerKa y tecnología que incorporan

¿Qué información
nos interesa?



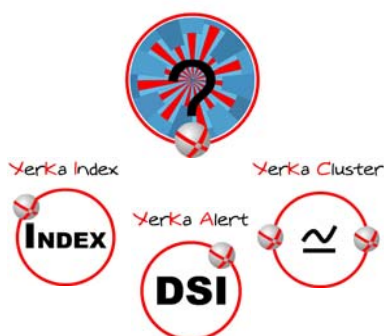
Espacio Temático

Encontrar la información
que nos interesa



Múltiples formas de
disponer la información

Rosa de los Temas



1 Módulo de gestión de espacios temáticos

Se encarga de la definición y realimentación del conjunto de temas que constituyen el espacio de interés de un usuario, grupo de trabajo u organización.

Los ejes temáticos se definen elaborando una terminología relevante en el tema. Estos conjuntos iniciales de términos se integran en la base de datos léxica utilizada por los analizadores lingüísticos así como en el sistema de argumentación de los metabuscadores del módulo de alimentación.

El resultado de las exploraciones se procesa por una serie de algoritmos orientados al "seguimiento terminológico de formas" capaces de recalibrar la relevancia de los términos componentes, detectar términos en desuso y descubrir el surgimiento de nuevos términos.

Los espacios temáticos dinámicos que se construyen son el armazón de los módulos de alimentación y recuperación del sistema XerKa.

2 Módulo de Alimentación

Se encarga de buscar y obtener los documentos que van a ser procesados por el Módulo de Análisis.

La búsqueda de sitios web se realiza generando consultas a buscadores de Internet con argumentos basados en los espacios temáticos definidos. El resultado de las búsquedas es preanalizado para valorar el interés dentro de cada espacio temático. Los argumentos de búsqueda se refinan haciendo cooperar en las búsquedas los términos que mejores resultados producen.

3 Módulo de Análisis

Efectúa diferentes análisis para nutrir al sistema de realimentación del gestor de espacios temáticos, al sistema de indexación del módulo de recuperación y al proceso de valoración de sitios Web del módulo de alimentación.

El contenido textual de los documentos se analiza en cuatro idiomas y se reduce a unidades de significado, sean lemas simples, lemas complejos o conceptos. Utilizando estas unidades en el contexto de los espacios temáticos definidos, se alimentan los algoritmos de seguimiento terminológico y los algoritmos de categorización.

La información estructural del documento y la información textual se utilizan para la tipificación o agrupación de documentos en tipos: cursos, curriculum, glosarios, etcétera.

4 Módulo de Recuperación

Se encarga de indexar los documentos en una base de datos relacional y multidimensional (MRDB), donde los documentos se reflejan bajo diferentes perspectivas orientadas a su Recuperación.

La perspectiva relacional de los índices temáticos facilita la recuperación de los documentos cuando la búsqueda se realiza bajo un punto de vista monotemático. La perspectiva multidimensional, sin embargo, responde inmediatamente a búsquedas por similitud dentro del espacio temático y produce también colecciones de documentos que puntúan simultáneamente en varios temas. Finalmente el **acceso por lema** nos permite efectuar búsquedas en la base de documentos establecida de forma similar a los buscadores de Internet ("XerKa - Index").

El acceso del usuario puede ser interactivo a través de una **interfaz de acceso multidimensional** ("Rosa de los Temas") o programado como **sistema de alertas** ("XerKa Alert - DSI")