



Themis
IUGWI2

THEMIS: Analizador y corrector de lenguaje sexista

MANUAL

Tipo Documento :	Manual
Versión:	4.0
Departamento :	Comercial
Fecha Creación :	11/06/2008
Última Modificación :	23/07/2009



ÍNDICE

1	CARACTERÍSTICAS DE LA HERRAMIENTA THEMIS	3
1.1	MÓDULO DE CONFIGURACIÓN Y GESTIÓN DEL DICCIONARIO	4
1.2	EDITOR DE ARCHIVOS	5
1.3	EXPLORADOR DE DIRECTORIOS	6
1.4	EXPLORADOR WEB.....	8
1.5	EXPLORADOR DE MAILS	8
1.6	DICCIONARIO THEMIS	9
2	OTROS ASPECTOS TÉCNICOS.....	10
2.1	INSTALADOR DE LA HERRAMIENTA	10
2.2	REQUISITOS MÍNIMOS DEL SISTEMA	10



1 CARACTERÍSTICAS DE LA HERRAMIENTA THEMIS

El lenguaje, como elemento que influye directamente en la concepción sociológica que se tiene de las personas, ha de ser cuidado en extremo para no caer en barbarismos ni exclusiones sociales. Es por ello por lo que desde The Reuse Company y otros muchos organismos, se está luchando por un uso no sexista del mismo, con la intención de alcanzar, en la medida de lo posible, un nivel de igualdad social entre mujeres y hombres.

Con el objetivo en mente de erradicar este lenguaje sexista de toda la documentación generada por cualquier institución, se pretende la utilización de herramientas informáticas que permitan explorar los documentos creados en busca de términos que puedan ser tachados de sexistas; consiguiendo así llegar a utilizar un lenguaje no sexista mediante la eliminación de la discriminación de género.

En este documento se detalla la funcionalidad de la herramienta Themis. Esta herramienta se ha desarrollado gracias a la perfecta conjunción entre un conjunto de lingüistas que han desarrollado un diccionario de entradas de contenido sexista; con más de 10 años de experiencia en el desarrollo de aplicaciones en el ámbito del Procesamiento del Lenguaje Natural.

La herramienta Themis se compone de cinco módulos independientes. Por un lado se dispone de una herramienta que, una vez instalada, permite explorar un documento (Word, Power Point, Writer, Impress, html, pdf...) en busca de terminología sexista y editar su contenido. El segundo módulo permite explorar directorios completos en busca de cualquier fuente ofimática (documentos textuales, generados con Word o Writer, PDF, presentaciones Power Point o Impress, o incluso páginas HTML) que contengan este tipo de términos. A través de un tercer módulo se posibilita el rastreo de páginas web de índole sexista, para lo que se comienza por la introducción de una URL (dirección Web) y el sistema rastreará todo el sitio web en busca de páginas que incluyan este tipo de terminología. Otro módulo Themis permite explorar nuestras carpetas de correo en busca de mails o adjuntos a mails que incluyan contenido calificado como sexista. Estos cuatro módulos son capaces de localizar documentos en los que aparezcan incluso los plurales de los términos incluidos en el diccionario de entradas sexistas.

Por último, la herramienta Themis incluye un módulo de administración que permite incluir las entradas al diccionario sexista que se consideren oportunas.

Themis – Herramienta proactiva e inteligente

Themis es una herramienta que podemos calificar de proactiva e inteligente:

- **Proactiva:** Themis NO es un diccionario de referencia, sino que de forma activa es capaz de rastrear documentación y sitios web en busca de aquellos términos que puedan estar siendo usados con carácter sexista
- **Inteligente:** gracias a Themis la aparición en diferentes lugares del documento de una misma palabra puede ser detectada, pero sólo se marcará allí donde su uso sea incorrecto



Themis marcará	Themis no marcará
Los mejores hosteleros del mundo...	Los principales centros hosteleros han tenido...
Los alumnos con calificación de ...	Los alumnos Juan y Pedro tienen calificación de...
Se precisan ingenieros informáticos	Se precisan ingenieros e ingenieras informáticos
Los españoles consumen...	Las y los españoles consumen...
Los atletas españoles han conseguido	Nuestras mejores atletas ya se han desplazado a...
El vecino	El vecino/a ¹ El vecino-a

A continuación se detallará pormenorizadamente la funcionalidad de cada uno de estos cinco módulos.

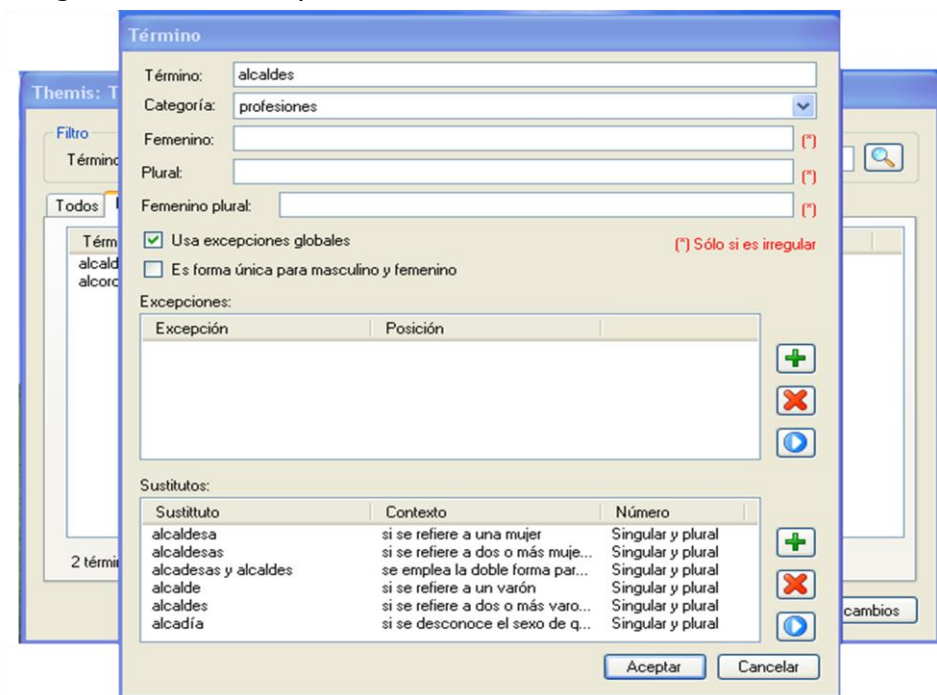
1.1 Módulo de configuración y gestión del diccionario

Se trata de una herramienta que permite, de modo muy simple, gestionar la base de datos donde reside el diccionario de entradas de índole sexista.

Por lo tanto, a través de esta herramienta las y los usuarios podrán:

- **Consultar entradas:** permitiendo filtrar la totalidad de las entradas y permitiendo entrar al detalle para cada una de ellas
- **Añadir/Modificar entradas**
- **Borrar entradas**
- **Filtrar entradas:** por tu texto

La siguiente figura muestra el aspecto de este módulo:



¹ Podrá marcarse o no en función de la configuración de la herramienta



1.2 Editor de archivos

Permite entrar en el contenido de un documento en particular y modificar su contenido para que se adecúe a las características de un documento libre de terminología sexista. Por este motivo, tras abrir el documento seleccionado, se ofrece la siguiente información:

- **Listado de términos localizados:** donde puede verse, uno por uno, todas las ocurrencias de términos localizados en el documento seleccionado
- **Contexto de aparición:** por cada término localizado puede verse el párrafo donde éste aparece, también se ofrece la posibilidad de ver un contexto más extendido, es decir, la totalidad del estado actual del documento
- **Sugerencias de cambio:** este listado muestra alternativas de uso válidas para el término marcado
- **Botones de actuación:**
 - **Omitir:** se genera una marca junto al término que indica que ha sido omitido por la usuaria
 - **Omitir todas:** se marcarán todas las apariciones del término como omitidas
 - **Cambiar:** cambia el término localizado por la sugerencia que el o la usuaria haya marcado
 - **Cambiar todas:** cambia todas las ocurrencias del término actual por la sugerencia que se haya seleccionado
 - **Cambiar por:** lo que ofrece la posibilidad de reemplazar el término seleccionado por aquel que el usuario o usuaria vea de mayor utilidad teniendo en cuenta el contexto de aparición del mismo

Una vez acometidos los cambios pertinentes, se puede volver a salvar el documento, con lo que se almacenarán estos cambios sobre el mismo. Los formatos de documentos que pueden tratarse con el editor son los siguientes:

- Microsoft Word: incluyendo documentos .doc, .docx, .dot, .dotx y .rtf
- Microsoft Power Point: incluyendo documentos .ppt, .pptx, .pot y .potx
- OpenOffice.org Writer
- OpenOffice.org Impress
- PDF
- Páginas web: incluyendo páginas almacenadas en discos duros locales o recursos de red en formato html y htm
- Texto plano: archivos con extensión .txt

El resultado del análisis puede exportarse como archivo html donde se incluyen:

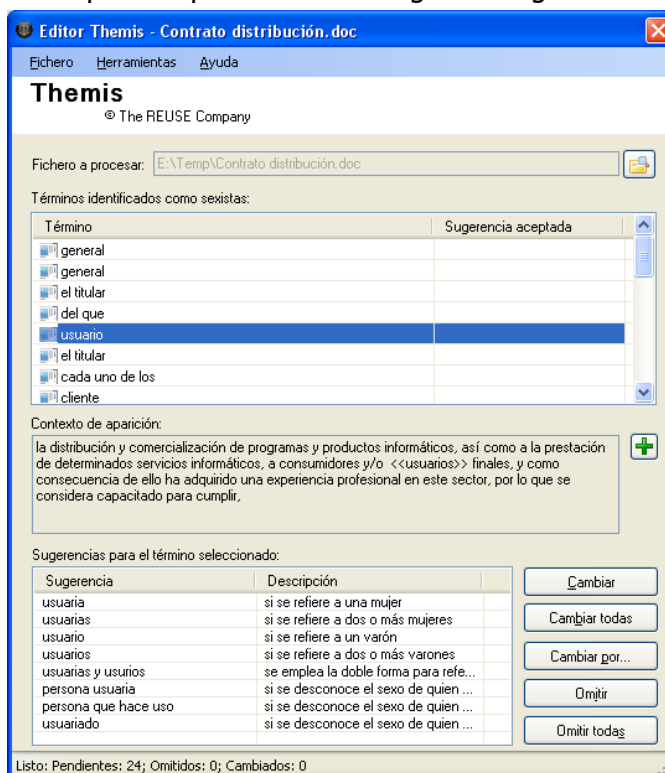
- Términos localizados
- Contexto de aparición
- Listado de sugerencias de cambio



El módulo editor de Themis puede arrancarse:

- Desde el menú de inicio de Windows
- Desde el explorador de Windows, haciendo clic con el botón derecho sobre uno de los archivos cuyas extensiones son reconocidas por el editor
- Sobre el Explorador de Archivos, haciendo clic con el botón derecho sobre uno de los archivos localizados

El aspecto de este módulo puede apreciarse en la siguiente figura:



1.3 Explorador de Directorios

Se trata de una herramienta con un aspecto similar a la herramienta de búsqueda de archivos que viene incluida dentro de los sistemas operativos Windows. La utilidad de esta herramienta es la de poder buscar documentos ofimáticos que contengan alguno de los términos de uso prohibido/restringido en las organizaciones.

Para ello, la usuaria podrá determinar:

- **El directorio donde comenzar la búsqueda:** esta búsqueda se efectuará desde la ruta indicada hasta cubrir la totalidad de archivos y subdirectorios ahí incluidos. La ruta puede teclearse en una caja de texto, seleccionar mediante un botón explorador o bien puede ser marcada desde el propio Explorador de Archivos de Windows haciendo clic con el botón derecho del ratón sobre un directorio
- **El nombre del documento:** permitiendo teclear un nombre concreto, o una secuencia de caracteres (incluyendo comodines) que determinen este nombre. P.e. "Propuestas*.doc"...



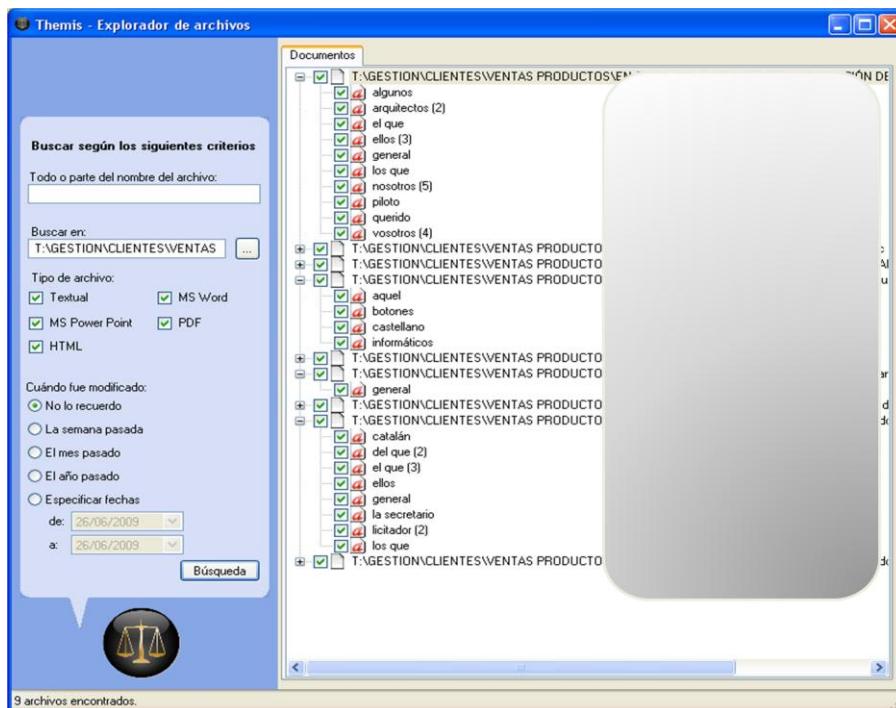
- **La tipología de los archivos rastreados:** permitiendo seleccionar alguna de las siguientes opciones: documentos textuales, documentos MS Word o OOO Writer, documentos PDF, presentaciones Power Point o OOO Impress o páginas HTML
- **La fecha de modificación de los documentos:** permitiendo determinar un rango de fechas, o bien buscar documentos modificados la última semana, el último mes o el último año

Según se van localizando archivos se van incluyendo en una lista donde, por cada archivo con contenido sexista, se incluyen tanto los términos localizados, como el número de ocurrencias de los mismos.

Tras finalizar el análisis, la usuaria o usuario tendrán opción de:

- **Generar un informe:** mostrando todos los documentos localizados, junto con los términos encontrados y el contexto donde aparecen
- **Editar el documento:** es decir, lanzar el módulo Editor de Archivos de Themis, lo que permitirá modificar su contenido para adecuarlo a las necesidades de un discurso libre de terminología sexista
- **Abrir el documento:** en este caso se abrirá la aplicación que el sistema operativo tenga vinculada con cada tipo de archivo

La siguiente figura muestra la apariencia de este módulo:





1.4 Explorador Web

Con un aspecto y modo de funcionamiento idéntico al Explorador de Directorios, permite rastrear sitios web completos en busca de cualquier página o documento colgado que pueda ser identificado como sexista.

Tras el análisis, el usuario o usuaria tendrán la opción de:

- **Generar un informe:** mostrando todos los documentos y páginas localizados, junto con los términos encontrados y el contexto donde aparecen
- **Abrir el documento:** lo que lanzará el navegador por defecto del sistema (p.e. Firefox, Internet Explorer...) con la página o documento marcado

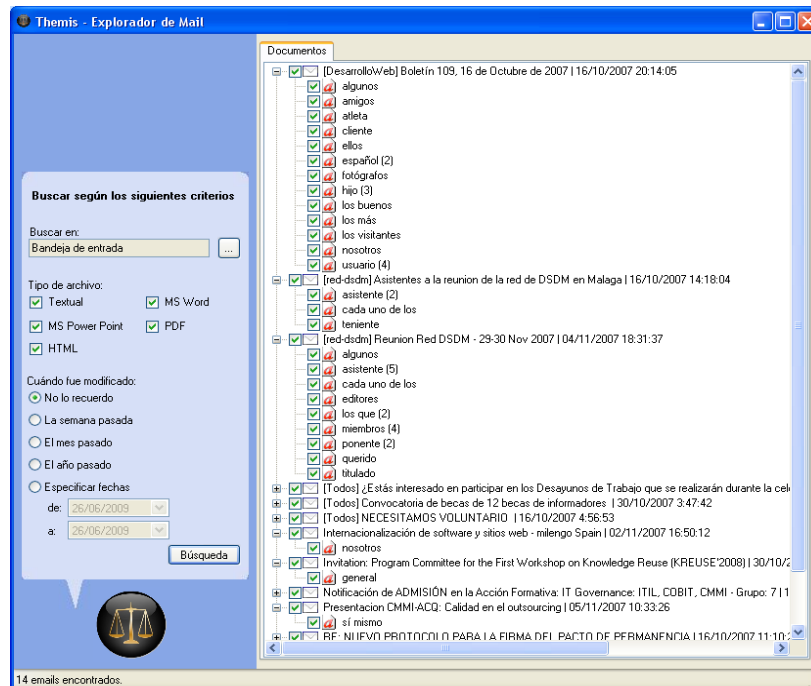
1.5 Explorador de Mails

De nuevo se trata de otro módulo con un aspecto y manejo idéntico al Explorador de Directorios. Este explorador de mails permite, partiendo de una carpeta de su correo electrónico, localizar aquellos documentos donde aparezcan términos sexistas. El explorador de mails es capaz de analizar, no sólo el cuerpo de los correos electrónicos, sino también el contenido de los documentos enviados como adjunto en estos mails².

Tras el análisis, el usuario o usuaria tendrán la opción de:

- **Generar un informe:** mostrando todos los mails localizados, junto con los términos encontrados y el contexto donde aparecen

El aspecto de este módulo puede verse en la siguiente imagen:



² Este módulo sólo es compatible con el gestor de correo Microsoft Outlook 2003 ó superior



1.6 Diccionario Themis

El núcleo de Themis es su completo diccionario. En la actualidad el diccionario de Themis incluye:

- Más de 1.600 términos
- Diferentes categorías a las que se asocia cada uno de los términos
- Más de 4.500 excepciones permiten que los diferentes términos no sean marcados siempre que aparecen en un documento
- Más de 16.000 sugerencias de cambio (una media de unas 10 excepciones por cada término)



2 OTROS ASPECTOS TÉCNICOS

2.1 *Instalador de la herramienta*

La instalación de la herramienta Themis se efectúa mediante un proceso de instalación convencional. Durante este proceso el o la usuaria simplemente deberá indicar el directorio de instalación (pudiendo dejar el directorio propuesto por defecto *C:\Archivos de Programa\Themis*).

Otros requisitos de la instalación, como puede ser la Máquina Virtual .Net v2.0 de Microsoft, son descargados e instalados automáticamente durante la instalación de Themis. Si el instalador verifica la existencia previa de este prerrequisito, este paso es automáticamente evitado.

2.2 *Requisitos mínimos del sistema*

Las herramientas incluidas en el presente contrato han sido desarrolladas utilizando la tecnología .Net de Microsoft. Por lo tanto, los requisitos mínimos son los siguientes:

- **Sistema operativos:** Microsoft Windows 2000, XP o superior
- **Paquete ofimático:** Microsoft Office 2000 ó Superior (probado exitosamente con Windows Vista y Office 2007)
- **Base de datos:** Microsoft Access. No siendo necesario que el puesto informático donde se instale la herramienta disponga de este gestor de bases de datos
- **Máquina virtual:** Máquina virtual de .net versión 2.0
- **Procesador:** Pentium III ó superior
- **Memoria:** 512 MB
- **Disco duro:** 30 MB de espacio libre

dTinf. The Reuse Company
Avda. Rey Juan Carlos I, 98, 4ºD
28916 Leganés, MADRID
SPAIN

| <http://www.themis.es/> | +34 91 680 90 22 | info@themis.es |

