

EXPERIENCIA EN LA IMPLANTACIÓN DE UN SISTEMA PARA GESTIÓN DE CONTENIDOS INTRANET E INTERNET EN EL INSTITUTO DE SALUD CARLOS III.

Fernando Martín Sánchez; María Dolores Marcos; Ana Zurdo;
Antonio Cárdenas; *Mercedes Alastruey; Luciano Sáez;
Unidad de Coordinación de Informática Sanitaria
**Unidad de Apoyo a la Dirección*
Instituto de Salud Carlos III

INTRODUCCIÓN

El Instituto de Salud Carlos III (ISCIII) es un organismo autónomo adscrito al Ministerio de Sanidad y Consumo que tiene como misión prestar soporte científico-técnico al Sistema Nacional de Salud. Entre sus principales competencias se encuentran: -epidemiología y salud pública, -evaluación de tecnologías sanitarias, -fomento y coordinación de la investigación biomédica y -docencia de postgrado. También desarrolla investigación en grupos de patologías (cáncer, cardiovasculares, enfermedades tropicales, enfermedades raras y anomalías congénitas) y presta servicios y asesoría técnica en nuevas tecnologías como la informática biomédica, la telemedicina o la genómica.

Para llevar a cabo sus funciones el ISCIII ha venido desarrollando desde el año 1996 diversos sistemas de información basados en tecnología Internet. (Intranet-Sirena, Museo Virtual de la Sanidad, página web pública). En los últimos tiempos existe una creciente cantidad y variedad de información que las distintas unidades del Instituto publican en Internet



y la demanda de este tipo de servicio es cada vez mas acusada. Por ello, se plantea la conveniencia de mejorar la gestión de contenidos en el seno de los sistemas I*net de los diversos centros.

En este sentido, para la Unidad de Coordinación de Informática Sanitaria, responsable de la provisión de servicios, educación e investigación en TIC para todo el Instituto, se considera contenido al conjunto de datos o información que se crean, archivan y publican en formato digital. Se entiende como publicada aquella información orientada a su consumo. Estos contenidos, por tanto, se pueden replicar, almacenar, distribuir y reutilizar y deben ser gestionados con eficiencia y seguridad.

Desde hace algunos años, en el mercado han venido apareciendo diversas soluciones lógicas comerciales y de libre uso específicamente orientadas a este objetivo. Los denominados SISTEMAS DE GESTION DE CONTENIDOS (SGC). Estos sistemas presentan como principales características que permiten separar el contenido de la presentación, lo que facilita cambios de diseño, aportan herramienta para descentralizar la publicación en Web (incluso para usuarios no-técnicos), incorporan módulos para gestión de un sitio web (workflow de creación, aprobación, publicación y archivo de contenidos), gestionan páginas generadas dinámicamente y posibilitan la personalización por usuario. Un aspecto clave en la gestión de contenidos es la categorización de la información, en base a la utilización de metadatos que permiten mejorar la gestión y la localización de información.

El circuito básico en los procesos de gestión de contenidos se describe a continuación en la figura 1.

De forma simplificada, hay que distinguir entre diversos roles en el proceso: -los autores (cualquier miembro de la organización que desee publicar un contenido), los publicadores (que revisan esa información y autorizan su publicación en tiempo y forma adecuadas) y los administradores, función técnica que consiste en optimizar el rendimiento y arquitectura del sistema, proponer las plantillas de diseño y los sistemas de categorización más adecuados, de acuerdo con los proveedores de información y de mantener el sistema en constante mejora y actualización. Finalmente los lectores acceden desde su navegador a los contenidos definitivos que la organización publique y pueden realizar búsquedas y recuperación de información.

Desde el punto de vista del SGC, todos los contenidos residen en una base de datos y el gestor responde a las solicitudes de páginas que se plantean desde los lectores recuperándolas desde aquella, componiendo las plantillas definidas y devolviendo al servidor web el contenido final que éste ofrece al lector.



CREAR Y GESTIONAR
(Gestión de elementos)

PRESENTAR
(Transformar)

ENTREGAR
(Publicar)

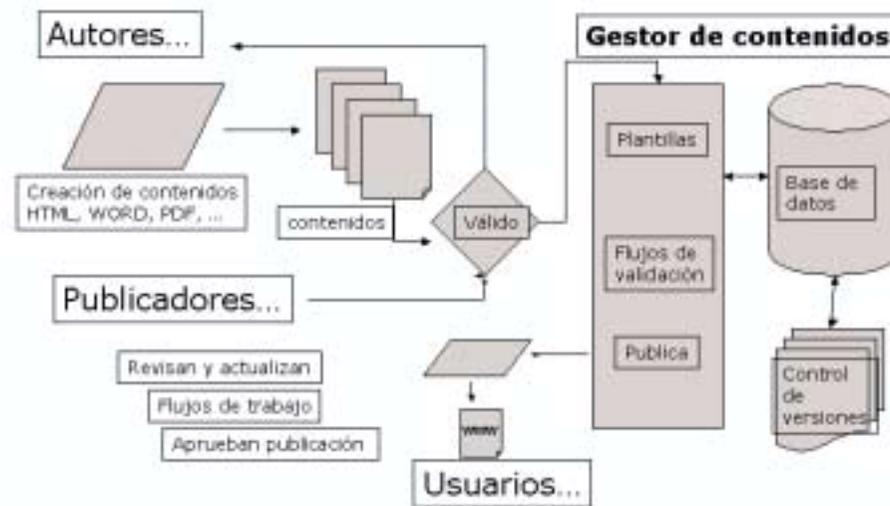


Figura 1. Esquema funcional de un sistema de gestión de contenidos (SGC)

Bajo estas premisas, en Noviembre de 2001 se plantea el proyecto de GC del ISCIII que se describe a continuación y que ha finalizado con la publicación de la nueva Intranet y portal web el día 15 de Julio pasado.



OBJETIVOS DEL PROYECTO

El proyecto de implantación de un sistema de gestión de contenidos para el portal Internet e Intranet en el Instituto de Salud Carlos III parte del deseo de la Dirección del Instituto de mejorar la situación existente a finales del año pasado en cuanto a prestación de servicios de información web internos y externos, y que se caracterizaba por:

- Existencia de varios servidores web no integrados dentro del portal corporativo <http://www.isciii.es> , en máquinas distintas y con URL diferentes. Aproximadamente se contaba con unas 2.500 páginas.
- Buscador no eficiente. Los índices no estaban suficientemente actualizados y en algunos servidores ni siquiera se disponía de la posibilidad de búsqueda. Además las páginas no estaban etiquetadas correctamente en su mayoría.
- Para las unidades funcionales del Instituto actualizar sus páginas no era un proceso sencillo, en muchos centros no existen expertos informáticos y los procesos de edición, validación y actualización se demoraban en el tiempo. Esto redundaba finalmente en la existencia de contenidos no actualizados.
- Por esta misma razón la publicación estaba muy centralizada en la Unidad de Informática, lo que de hecho llegaba a significar un cuello de botella en el proceso.
- La navegación por la página web principal era monolítica, totalmente basada en el concepto de organigrama, es decir el usuario tenía que conocer la estructura y las competencias de las unidades y subdirecciones del Instituto para llegar a encontrar la información deseada en muchos casos (navegación por unidades).
- Finalmente, era impensable realizar cualquier tipo de personalización de contenidos en función de distintos perfiles o intereses de los usuarios.

Por tanto, los objetivos del proyecto pasaban por dar una respuesta tecnológica a estos problemas, en un tiempo breve (6 meses), con un presupuesto moderado, con recursos humanos limitados y con la premisa de que debía mantenerse la imagen corporativa y por tanto el diseño básico del web existente, así como asegurar la incorporación de los contenidos ya existentes, es decir, no se debía repetir la escritura de ninguna de las páginas ya publicadas.



METODOLOGIA Y GESTIÓN DEL PROYECTO

El proyecto se gestionó desde la Unidad de Coordinación de Informática Sanitaria y la Unidad de Apoyo a la Dirección, atendiendo a 6 áreas de desarrollo:

1. Selección de la herramienta:

Se eligió i.Reach de Ascential Software por ser uno de los pocos SGC por entonces incluidos en el Catalogo de Bienes Centralizados de Patrimonio y el único que encajaba en los condicionantes de presupuesto y rapidez de aprendizaje e implantación, tras la realización de una evaluación, seguida de una prueba y de una maqueta del sistema que resultó adecuada para las funcionalidades deseadas. El aspecto básico se muestra en la figura 2.



Figura 2. Pantalla del gestor de contenidos



2. Formación y grupo de usuarios:

Se diseñó un programa que consistió en la realización de cursos “on-site” sobre los siguientes aspectos:

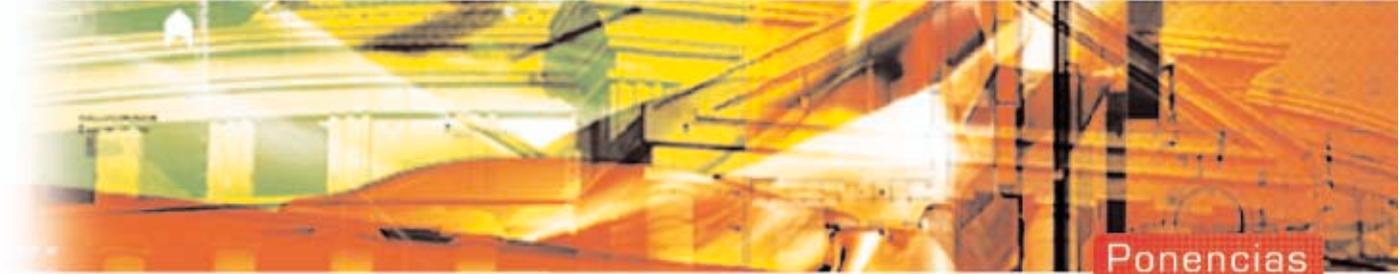
- 2 cursos de usuario
 - Publicación (1 mañana) : 22 personas, 2 sesiones
 - Administración (1 mañana) : 9 personas, 1 sesión
- Programación (evaluándose prerequisites: HTML, Javascript, SQL, SPL)
 - Herramienta programación: 4 días
 - Producto: 3 días
- Administración Base de Datos: 5 días.

A estos módulos, dirigidos a técnicos y a publicadores o revisores, se les sumó la programación de una serie de cursos de usuario para autores en el Plan de Formación Interna del Instituto, impartidos por el propio personal de la Unidad TIC. Hay que decir que a lo largo de todo el proceso se contó con el apoyo de la Dirección, que desde el inicio constituyó una “Comisión web” en la que estaban representados todas las unidades del Instituto (unas 20 personas) y que se fue reuniendo periódicamente para ser informada y comentar las diversas incidencias a través de una lista de correo electrónico gestionada desde la Unidad de Apoyo a la Dirección.

3. Instalación, arquitectura y seguridad

En la figura 3 se representa esquemáticamente la arquitectura que resultó finalmente elegida tras valorar diversas opciones entre la empresa suministradora y el Departamento de Sistemas del Instituto.

Los aspectos más destacables serían la instalación de una única instancia del gestor de contenidos en una máquina biprocesador dedicada, con dos bases de datos (Informix) una para contenidos Intranet y otra para contenidos Internet, la decisión de realizar links desde Intranet a contenidos web cuando sea pertinente y la creación de una parte pública,



de posible acceso mediante registro y de una parte privada. El registro pues, consiste en que un usuario personaliza su presentación seleccionando su perfil y cuando accede con su login y contraseña tiene acceso a contenidos no visibles para el resto de los usuarios.

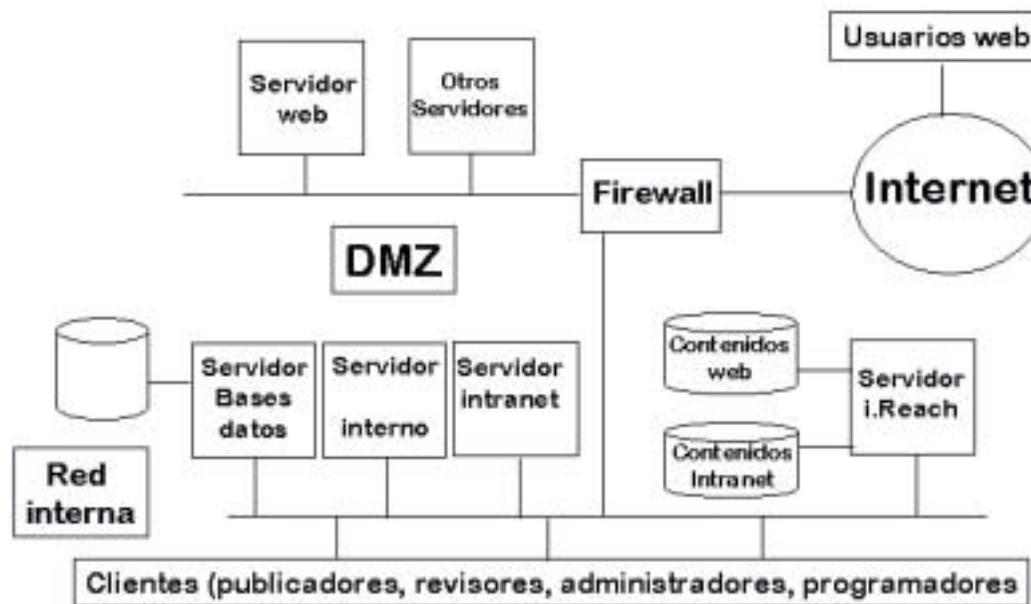


Figura 3. Arquitectura tecnológica



4. Migración y diseño

Como se comentó anteriormente, el proyecto no consistía en la realización de un web “de novo”, sino en la integración en el nuevo sistema de GC de las páginas existentes. Por ello, uno de los grupos de trabajo técnicos se orientó hacia la elaboración de herramientas que facilitarían la migración de contenidos hacia el sistema nuevo y en asistir a las distintas unidades funcionales en la realización de este proceso. Realmente sin la colaboración de todas ellas no se habrían podido alcanzar los plazos previstos. Este mismo grupo diseñó también, con la ayuda de la empresa suministradora, las plantillas básicas, los menús y las composiciones del nuevo sistema. Además se creó un pie común para todas las páginas y un “track” automático, para guía de la navegación. En algún momento fue necesaria una labor de “limpieza” de código html de páginas existentes y en todos los casos, los propios usuarios procedieron a categorizar las páginas atendiendo a los criterios y atributos que se describen en la siguiente sección.

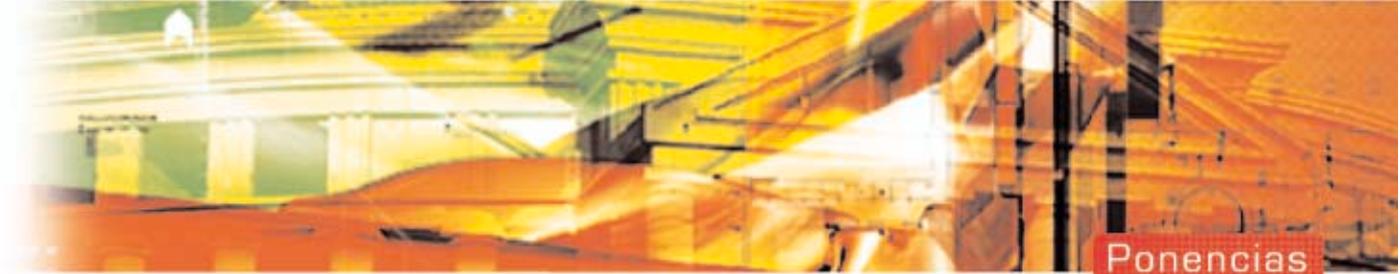
Por último, se decidió, por parte de la dirección del proyecto, habilitar un nuevo modo de navegación tanto del Web como de la Intranet basado en el concepto de actividad, adicional al ya existente de navegación por unidad. Para ello fue necesario que los autores de las páginas realizaran “shortcuts” o enlaces virtuales desde cada una de sus páginas al lugar apropiado de la nueva estructura.

5. Categorización y búsquedas

Realmente el diseño de un buen sistema de categorización de contenidos resulta una actividad central en un proyecto de este tipo. La correcta creación y asignación de atributos a los contenidos es un aspecto clave para optimizar la navegación, la localización y recuperación de información pertinente.

En este proyecto se diseñó un sistema según el cual, como se describe en las figuras 4 y 5, cada contenido forma parte de un espacio de 4 dimensiones o atributos; podría decirse por tanto que cada página lleva asignado un vector de 4 coordenadas (uno especial: -tipo y 3 categorías: -unidad, -actividad y -área temática). De este modo, para cada ítem, queda perfectamente definido:

- QUE ES (tipo),
- DE QUIEN (unidad funcional responsable),
- PARA QUE (actividad) y
- SOBRE QUE (área temática).



De estos 4 atributos, unidad y tipo son compartidos para Intranet-Internet, mientras que actividades y áreas temáticas son distintas en estos dos entornos.

Este sistema permite la deseada navegación por actividades, aunque exige, como se ha comentado anteriormente la categorización retroactiva de páginas ya existentes y la elaboración de normas concretas para la categorización de nuevas páginas. Para ello se elaboró y distribuyó un pequeño manual en el que se recogía un árbol inicial de atributos con sus valores posibles, que se fue enriqueciendo y refinando a lo largo del proyecto.

Por ejemplo, algunos valores aceptables para el atributo especial TIPO son: -nota de prensa, -noticia, -convocatoria, -informe, -procedimiento o -aplicación. Para el atributo UNIDAD se recoge el organigrama del Instituto hasta en 3 niveles (subdirección, centro, grupo). Para el atributo ACTIVIDAD el primer nivel incluye los valores -servicios, -investigación y -docencia y en el segundo nivel se detallan los tipos de cursos, los distintos servicios que presta el Instituto o distintos aspectos relacionados con la investigación (patentes, proyectos). Por último, el árbol de AREAS TEMÁTICAS recoge en 2 niveles los campos de conocimiento de interés para el personal del Instituto (salud pública, microbiología, epidemiología, informática, ...).

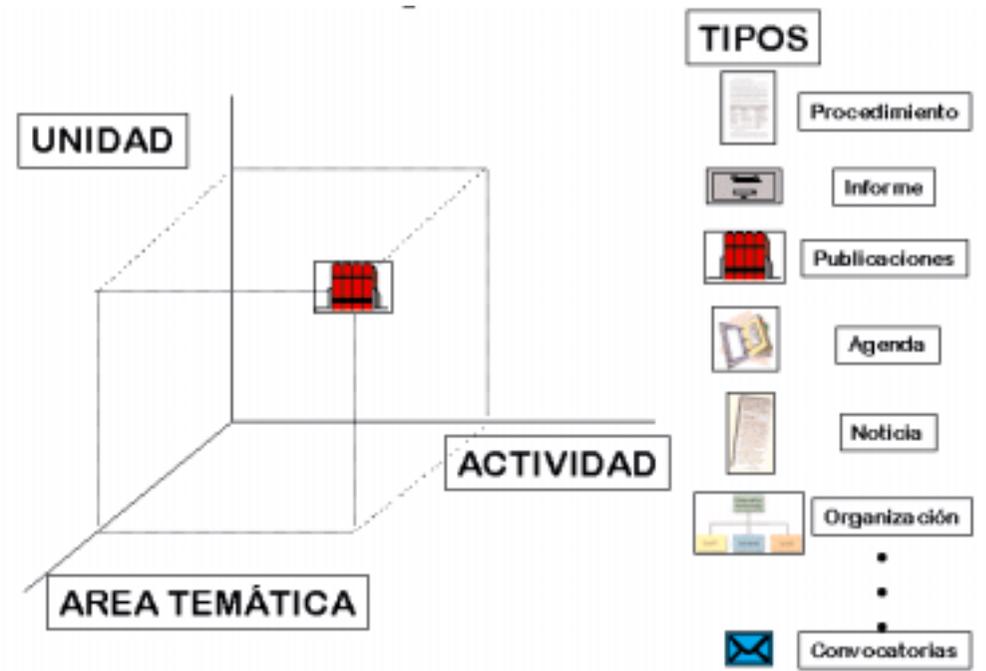
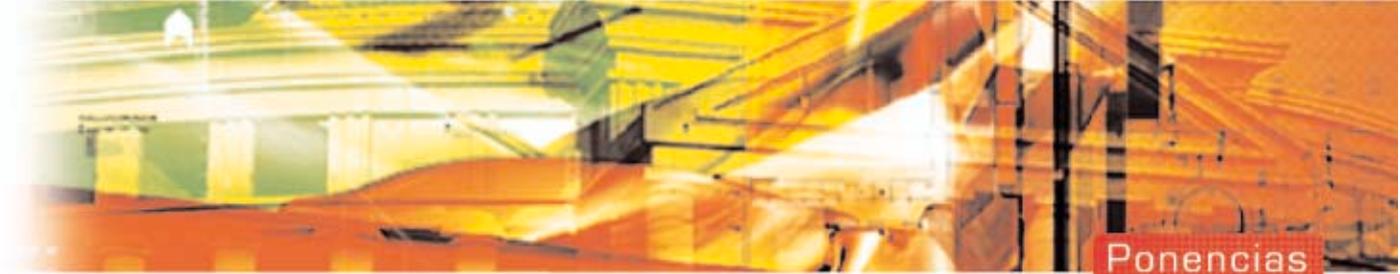


Figura 4. Sistema de categorización de contenidos



Este sistema de categorización, que se ilustra con un ejemplo en la figura 5, ha permitido explotar al máximo el motor de búsqueda incorporado en el SGC (basado en Excalibur), ofreciendo a nuestros usuarios un completo buscador que puede trabajar en modo simple (campo texto sobre título y resumen) o en modo avanzado (búsqueda en contenido de páginas por atributos desplegados) en 3 pestañas correspondientes a los atributos (tipo de contenido, unidad, actividad, área temática), con restricciones de fechas, uso de operadores lógicos y elección de rangos de similitud (búsqueda difusa, o exacta).

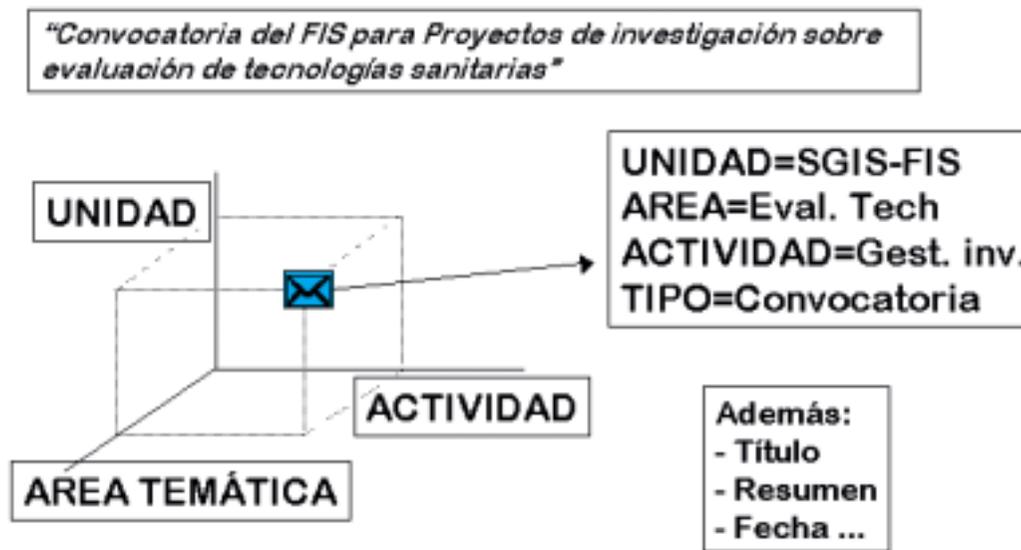


Figura 5. Un ejemplo de categorización de un documento público



6. Flujos de validación

El último aspecto del proyecto consistió en la definición de los circuitos de validación y aprobación de documentos para su publicación, representados en la figura 6. Una vez estudiada la estructura de las diversas unidades del Instituto se decidió acometer este aspecto en dos fases, por razones de agilidad y eficiencia. En una primera fase existirían 2 niveles: -1 supervisor de contenidos (la dirección) y 17 publicadores (uno por unidad funcional). De este modo solo un número reducido de personas por unidad tenían privilegios para publicar las paginas de su ámbito. En la segunda fase, que se está implementando en la actualidad, se amplía a 3 niveles: 1 supervisor (dirección), 17 revisores o publicadores y N publicadores (siendo deseable que este número N llegue a ser el total de personas que pueden ser autores de algún tipo de información o datos para publicar).

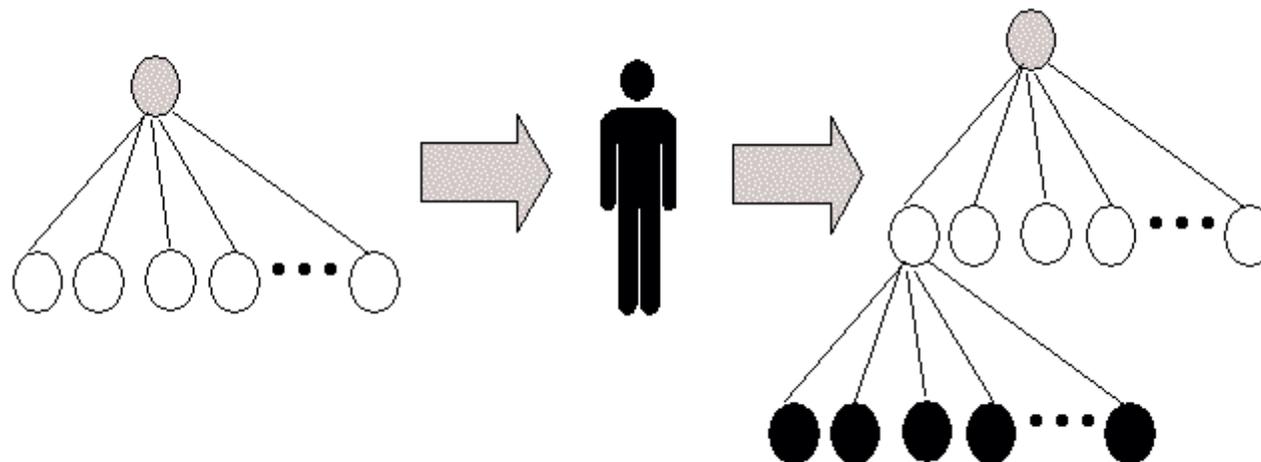
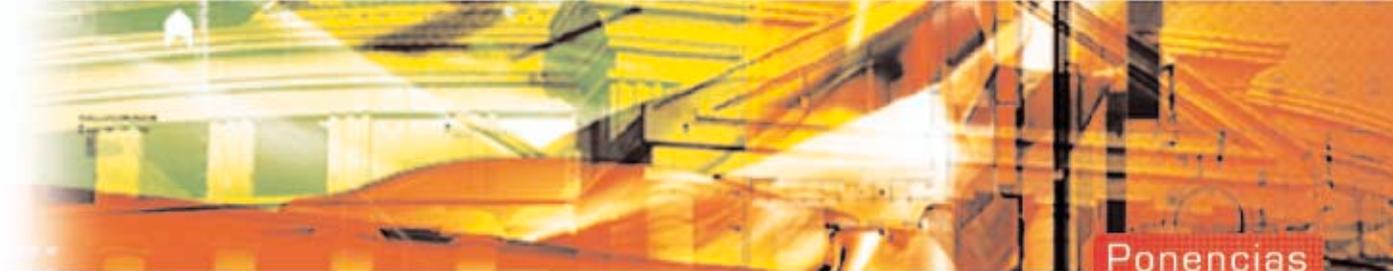


Figura 6. Fases del proyecto en los circuitos de validación



RESULTADOS

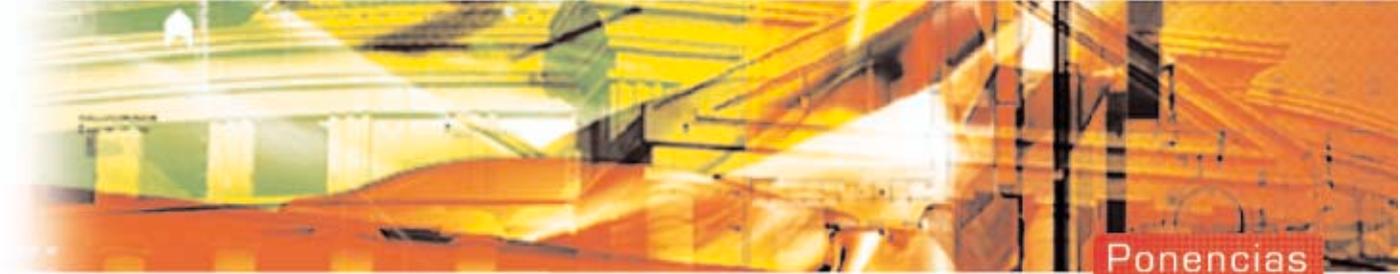
El portal Intranet completo se publicó en el mes de marzo y el portal Internet es accesible en <http://www.isciii.es> desde el pasado día 15 de Julio (Figura 7). Se han migrado al nuevo sistema más de 2.000 páginas y se han integrado 4 de las 6 URL independientes que existían. Las dos restantes tienen prevista su inclusión en el nuevo sistema a lo largo de este mismo año.

Previamente a la publicación de ambos sistemas se estableció un pequeño periodo de pruebas para cada uno (una semana) y se presentó el resultado final tanto a la Comisión web como a la Comisión de Dirección.

Se ha completado la documentación técnica del sistema y los manuales de uso y administración y sólo quedan pendientes algunos cursos de usuario y la elaboración de nuevas plantillas y diseños para incorporar sugerencias principalmente estéticas provenientes de los usuarios.



Figura 7. Portal Internet del ISCIII



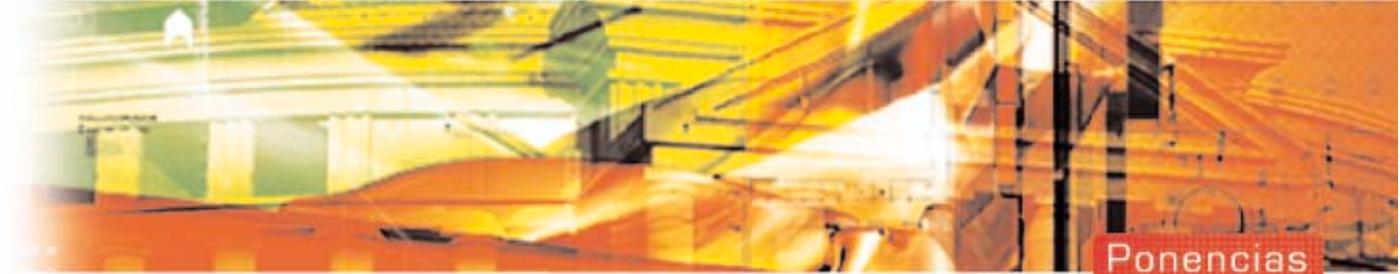
DISCUSION

Los principales problemas encontrados han consistido en la dificultad de introducir un nuevo modo de trabajo descentralizado de administración de las páginas web de las unidades, en alcanzar acuerdos sobre la mejor organización de la información en aquellas áreas en las que varias unidades tienen competencias relacionadas y en consolidar los circuitos de validación de los nuevos contenidos.

Entre los beneficios obtenidos habría que destacar la homogeneidad y consistencia del interfaz de usuario, la información previa sobre los documentos (ver figura 8), el sistema de vigencia de páginas, el rendimiento y estabilidad de la instalación, el sistema de categorización y el buscador avanzado. Otros aspectos positivos han sido el sistema de personalización, la integración de otros URLs, las ayudas a la navegación (track y mapa automático), las estadísticas y contadores, el acceso a aplicaciones, el doble modo de navegación por unidades y por actividades, la descentralización de la publicación y la gestión de versiones en inglés de las páginas existentes en ambos idiomas.



Figura 8.
Una página con carpetas y documentos del portal Internet del ISCIII



FUTUROS DESARROLLOS

La implantación del SGC en la Intranet y el portal Internet del ISCIII va a permitir en el futuro una fácil generación de contenidos para otros canales (PDAs, WAP) y la posibilidad de sindicación y agregación de los mismos con otras entidades colaboradoras.

También se deberá avanzar en aspectos solamente iniciados con esta fase del proyecto como son -la personalización, -la integración con aplicaciones o -el multilingüismo.

Por último, el sistema nos abre las puertas de nuevos proyectos en los que se asocie el SGC con sistemas de gestión de conocimiento y de e-learning que ya se han puesto en marcha en nuestro centro. Con este trabajo se ha dado el primer paso para preparar a los sistemas de información Internet del Instituto para la que seguramente constituirá la nueva revolución en Internet: -la web semántica.