

## 5. Área de informática

### 5.1. Proyectos ejecutados

#### 5.1.1. Proyectos de apoyo a la gestión

- 5.1.1.1. Plan de Sistemas de Información
- 5.1.1.2. Gestión de proyectos
- 5.1.1.3. Gestión de LOPD

#### 5.1.2. Proyectos de infraestructura

- 5.1.2.1. Adecuación del Centro de Proceso de Datos del ICANE
- 5.1.2.2. Mejoras en infraestructura de virtualización y almacenamiento centralizado
- 5.1.2.3. Migraciones y actualizaciones de hardware/software en sistemas críticos
- 5.1.2.4. Renovación de infraestructura de red perimetral

#### 5.1.3. Nuevas implantaciones de sistemas

- 5.1.3.1. Suite de mensajería y colaboración
- 5.1.3.2. Arquitectura de integración de datos basada en la *Suite Pentaho*
- 5.1.3.3. Motor documental

#### 5.1.4. Proyectos de explotación de datos estadísticos

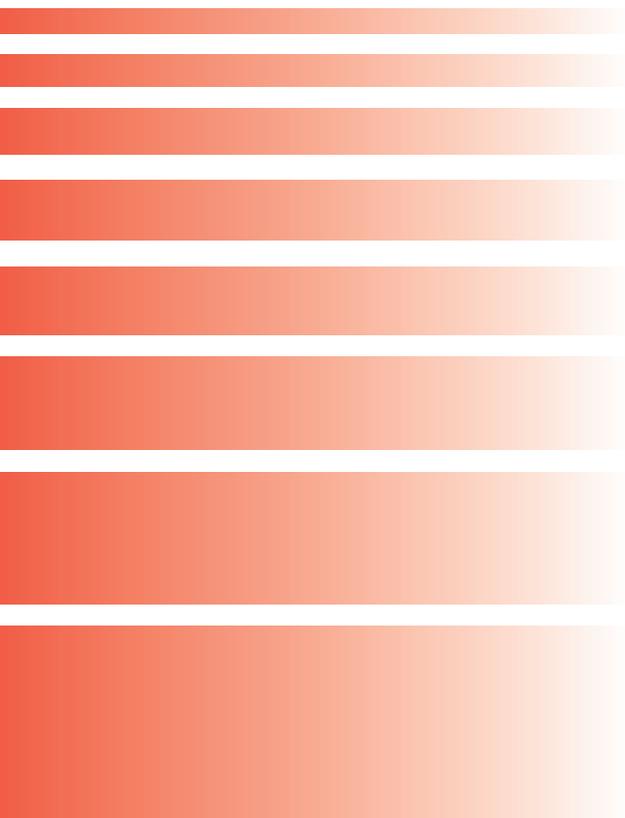
- 5.1.4.1. Encuesta Marco Input-Output 2007
- 5.1.4.2. Encuesta Social de Cantabria 2009
- 5.1.4.3. Estadística de Comercio Exterior
- 5.1.4.4. Encuesta de Ocupación Hotelera
- 5.1.4.5. Movimiento Natural de Población
- 5.1.4.6. Fichas Municipales 2008
- 5.1.4.7. Directorio de Empresas 2008

## **5.2. Proyectos en marcha**

5.2.1. Reestructuración del sitio web en un portal corporativo

5.2.2. Reestructuración del banco de datos

5.2.3. Automatización del sistema de publicaciones





## 5.1 PROYECTOS EJECUTADOS

### 5.1.1. PROYECTOS DE APOYO A LA GESTIÓN

#### 5.1.1.1. Plan de Sistemas de Información

El Plan de Sistemas de Información (en adelante PSI) del ICANE tiene como objetivo la obtención de un marco de referencia para el desarrollo e implantación de sistemas de información que responda a los objetivos estratégicos de todas las áreas del Instituto.

Para la elaboración del PSI se han estudiado las necesidades de información de los procesos afectados por el mismo, con el fin de definir los requisitos generales y obtener modelos conceptuales de información. Por otra parte, se han evaluado las opciones tecnológicas, proponiendo un entorno adecuado a las mismas.

Tras analizar las prioridades relacionadas con las distintas variables que afectan a los sistemas de información, se ha elaborado un calendario de proyectos con una planificación lo más detallada posible de los más inmediatos. Además, se ha propuesto un procedimiento para mantener actualizado el PSI para incluir en él todos los cambios necesarios, garantizando el cumplimiento adecuado del mismo.

El listado de proyectos confeccionado para 2009 por los motivos expuestos ha sido el siguiente:

- Mejora de los flujos de información y procesos colaborativos entre los Servicios del ICANE, y dentro de los mismos.
- Mejora del sistema de actualización del sitio web del ICANE, ofreciendo alternativas para subsanar errores de manera que puedan ser utilizadas por personal no técnico. Establecimiento de un sistema interactivo de carga de datos procedentes de Internet y de ficheros sencillos, o directamente de publicaciones recibidas o descargables en formatos estándares.
- Integración, replicación o unificación de los motores de bases de datos público o externo y privado o interno.
- Evaluación e incorporación de herramientas o sistemas para mejora de los procedimientos de análisis, consulta, integración y calidad de datos.
- Mejora y aumento del porcentaje de automatización de los procesos de producción de publicaciones.
- Reestructuración total del banco de datos incorporando notas metodológicas, mejorando las fuentes de las series estadísticas, aumentando la organización de la clasificación temática, mejorando su facilidad de uso, ampliando las posibilidades de cruce y optimizando el número de series a incluir en su banco de datos. Análisis de las consultas realizadas y reducción el número de series almacenadas, incluyendo las que tienen demanda y las que se precisen para otras aplicaciones.

## ÁREA DE INFORMÁTICA

El seguimiento del PSI se ha realizado a través de reuniones mensuales o bimestrales, en función de la carga de trabajo y la disponibilidad de todo el personal.

### 5.1.1.2. Gestión de proyectos

Con el fin de mejorar los diversos canales de comunicación entre servicios y dentro de los propios servicios existentes en el ICANE, se ha implantado un sistema de gestión de proyectos multi-usuario que permite una organización más eficaz. Se trata de una aplicación web de fuentes abiertas basada en tecnología PHP que permite, entre otras, las siguientes funcionalidades:

- Descomposición de los diversos proyectos ejecutados por el ICANE en sus distintas tareas, estableciendo hitos, fechas, etc.
- Asignación de personal participante a las distintas tareas y envío de notificaciones.
- Producción de cronogramas (gráficos de Gantt).
- Inclusión de documentación electrónica asociada a cada proyecto.

El sistema descrito permite a los técnicos del ICANE la creación de un repositorio o catálogo de proyectos que sirva de referencia para el desarrollo y mejora de futuros trabajos.

### 5.1.1.3. Gestión de LOPD

Dada la naturaleza de determinados conjuntos de datos alojados en la infraestructura del ICANE, ha sido imperativa la adaptación de los distintos procedimientos implicados para dar cumplimiento a la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal y al reglamento que la desarrolla (Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal).

Para ello, se ha contratado un modelo de aplicación de gestión basada en SaaS (Software as a Service), "GESDATOS", que permite dar soporte a las tareas relacionadas con el cumplimiento de la ley, destacándose entre sus características las siguientes:

- Desarrollo de diagnósticos preliminares sobre el estado del Instituto y las políticas a desarrollar.
- Generación automática del Documento de Seguridad.
- Gestión cotidiana y automatizada de las políticas de seguridad.
- Establecimiento de un sistema de alertas sobre tareas pendientes.
- Desarrollo de controles periódicos y auditorías.

El Centro de Procesos de Datos (en adelante CPD), donde está instalada la aplicación, cuenta con sendas certificaciones ISO 9001 e ISO 27001, siendo el SLA (Acuerdo de Nivel de Servicio) para la disponibilidad del servidor que lo aloja de al menos un 98,8% en un mes natural dado.



## ÁREA DE INFORMÁTICA

Del mismo modo cuenta con distintas garantías de climatización, suministro ininterrumpido de corriente, controles de acceso físico, detección y extinción de incendios y monitorización 24x7.

Asimismo, con objeto de mejorar las tareas exigidas por la ley, se han realizado las siguientes mejoras en el sistema de copias de seguridad en cinta:

- Se añaden los grupos de cintas de rotación mensual/anual.
- Se reestructuran y simplifican las tareas de copia para facilitar la gestión, pasando de 20 a 3.
- Se pasa a utilizar dos cintas en lugar de una para las copias completas.
- Se activa la multiplexación en las tareas, permitiendo que se haga copia de varias fuentes distintas al mismo tiempo, lo que resulta tanto en un tiempo de copia menor (se pasa de una media de 400 MB/s a 1200 MB/s, con lo que una copia completa requiere unas 6h) como en una prolongación de la vida de las cintas.
- Se implementa un calendario compartido con la información de las rotaciones.

### 5.1.2. PROYECTOS DE INFRAESTRUCTURA

#### 5.1.2.1. Adecuación del Centro de Proceso de Datos del ICANE

Se ha adecuando una nueva estancia para ubicar el CPD en la 2ª planta del edificio donde se ubican las oficinas del ICANE, incorporando las siguientes mejoras y/o trabajos realizados:

- Revisión y certificación de cableado estructurado.
- Dotación de mayor espacio físico para crecimiento futuro.
- Puesta en marcha de dos unidades de ventilación en modo activo/pasivo.
- Inclusión de dos unidades de alimentación ininterrumpida adicionales con capacidad de 3 kW cada una y conexión en red.
- Dos nuevos conmutadores de red a Gigabit, conectados a través de troncales @1Gbps. Creación de agregado @2Gbps para conexión de la infraestructura de virtualización.
- Incorporación de un control de accesos basado en RFID con baterías para funcionamiento autónomo que permite gestionar hasta cuatro puertas, registrar las entradas de los usuarios en las oficinas y abrir cerraduras motorizadas.

Los trabajos de mejora permitirán el cumplimiento de la Ley Orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal y el reglamento que la desarrolla (Real Decreto 1720/2007, de 21 de diciembre, por el que se aprueba el Reglamento de desarrollo de la Ley Orgánica 15/1999, de 13 de diciembre, de protección de datos de carácter personal. Del mismo modo, pretenden acercar el CPD a las recomendaciones expuestas en la norma TIA/EIA 942, que si bien no es vinculante, ofrece un conjunto de buenas prácticas relacionadas.

## ÁREA DE INFORMÁTICA

### 5.1.2.2. Mejoras en infraestructura de virtualización y almacenamiento centralizado

A través de la adquisición e implantación de una Red de Área de Almacenamiento (SAN, Storage Area Network) basada en tecnología iSCSI, se han podido obtener las siguientes ventajas:

- Simplificación, facilitación y optimización de las tareas de provisión y administración (parcheado de seguridad, actualización, etc.) de servidores físicos y virtuales.
- Realización de copias de seguridad de primer nivel tanto en frío como en caliente, a través de un medio seguro, fiable, modular y escalable con el tiempo.
- Creación de catálogos de instantáneas o snapshots de máquinas virtuales para recuperación total casi inmediata en caso de contingencia.
- Activación de un sistema de alta disponibilidad capaz de ejecutar las máquinas virtuales en otros servidores físicos en caso de fallo.

La unidad de almacenamiento ha sido montada en configuración RAID-5 y con dos controladoras independientes en modo activo/pasivo, utilizando un direccionamiento de red privado y único al efecto.

### 5.1.2.3. Migraciones y actualizaciones de hardware/software en sistemas críticos

Los siguientes sistemas o servicios han sido virtualizados y actualizados a nuevas distribuciones y versiones:

- Servicio web y de correo:

	Situación anterior	Situación actual
<b>CPU</b>	Intel Xeon @ 3.2 GHz	2 núcleos virtuales Intel Xeon QuadCore E5420 @2,50 GHz
<b>RAM</b>	1 GB (PC2100 DDR SDRAM)	4 GB PC2-5300 FBD DDR2-667 SDRAM
<b>Conexión de red</b>	Ethernet @100 MBps	Compartido Ethernet @1Gbps
<b>Disco</b>	DAS Wide Ultra 3 SCSI	SAN iSCSI @ 1Gbps, SAS
<b>Sistema operativo</b>	RHEL 3 - 32 bits	RHEL 5.4 - 32 bits

- Sistema gestor de bases de datos interno:

	Situación anterior	Situación actual
<b>CPU</b>	Intel Xeon @ 3.2 GHz	4 núcleos virtuales Intel Xeon QuadCore E5420 @2,50 GHz
<b>RAM</b>	1 GB (PC2100 DDR SDRAM)	4 GB PC2-5300 FBD DDR2-667 SDRAM
<b>Conexión de red</b>	Ethernet @100 MBps	Compartido Ethernet @1Gbps
<b>Disco</b>	DAS Wide Ultra 3 SCSI	SAN iSCSI @ 1Gbps, SAS
<b>Sistema operativo</b>	RHEL 3 - 32 bits	RHEL 5.4 - 64 bits
<b>Sistema gestor de bases de datos</b>	Oracle 10g Release 1	Oracle 11g Release 2

Ambas máquinas se encuentran ahora desplegadas en un entorno virtual de alta disponibilidad. El almacén de datos ha sido migrado completamente, reestructurando la organización interna del mismo para adecuarlo a la nueva versión del motor.



#### 5.1.2.4. Renovación de infraestructura de red perimetral

El valor creciente de la información y de los sistemas de tecnologías de la información que la soportan, su omnipresencia y su carácter de instrumento esencial para el desarrollo económico y social de nuestra sociedad, las dependencias que se dan, y los riesgos generados, conducen todos ellos a la necesidad de adoptar políticas, procedimientos, prácticas y medidas organizativas y técnicas capaces de proteger la información y de gestionar la seguridad de los sistemas respondiendo a las amenazas existentes; capaces de garantizar dimensiones esenciales de la seguridad como la confidencialidad, la integridad, la disponibilidad y la autenticidad; de satisfacer la confianza depositada en los productos y sistemas, en la información necesaria para la toma de decisiones y en las posibles expectativas en cuanto a oportunidades de innovación y adaptación; así como de satisfacer los posibles requisitos legales, sean éstos de carácter horizontal o sectorial.

La dependencia de los sistemas de información preocupa cada vez más a la sociedad ya que genera riesgos debidos a la propia complejidad de los sistemas, a posibles accidentes, errores o ataques, a la constante evolución en un entorno cambiante, o a un posible uso irresponsable de los mismos. La materialización de estos riesgos puede afectar a la propia continuidad de los servicios (internos y externos), a la protección de la información en general y, en particular, de los datos de carácter personal, así como a la propia validez y eficacia de los actos que se apoyan en transacciones electrónicas, por ejemplo de administración o comercio electrónico.

Los diversos actores afectados, particulares, administraciones públicas y empresas, reclaman seguridad y, en definitiva, confianza en el uso de los sistemas de tecnologías de la información.

Los pasos a dar para garantizar la seguridad de los sistemas de tecnologías de la información se orientan a la implantación de una gestión continua de la seguridad, a la adopción de controles y salvaguardas organizativas y técnicas que garanticen aspectos tales como la continuidad de su funcionamiento, la protección de la información, la validez de las transacciones electrónicas, la conformidad con el marco normativo y contractual correspondiente, con condiciones tecnológicas (estándares) determinadas, el aseguramiento en cuanto a un uso adecuado y optimizado de los recursos, y, en general, la satisfacción de aquellos requisitos que contribuyen al logro de los objetivos de la organización.

El ICANE, como organismo público, no es ajeno a estos hechos y se encuentra en continua y profunda revisión de sus mecanismos de control y salvaguardas tanto organizativas como técnicas. En este último contexto se desarrolla este proyecto, que pretende solucionar una serie de carencias tecnológicas y de seguridad perimetral arrastradas en el tiempo, así como añadir nuevas funcionalidades y mecanismos de control que permitan al Servicio de Informática y Banco de Datos facilitar labores de administración y mantenimiento.

La renovación de la infraestructura de red perimetral consistió en la sustitución de un enrutador y un cortafuegos obsoletos por dos más adecuados a las necesidades del momento actual, cumpliendo con las siguientes características:

- Uso de tecnologías y mecanismos de seguridad basados en estándares y en conformidad con el estado del arte en materia de seguridad de sistemas de información, entendiendo como requisitos mínimos el filtrado de paquetes con y sin

## ÁREA DE INFORMÁTICA

estado, mecanismos de detección y prevención de intrusiones, filtrado de contenidos y anti-virus y anti-malware en general.

- Incorporación, por parte del cortafuegos, enrutador o ambos de tecnologías de creación de túneles VPN IPSEC, y en especial de la tecnología SSL-VPN, dada la naturaleza web de los sistemas y aplicaciones existentes en el ICANE.
- Toda la electrónica de red cuenta con una consola de gestión o administración sencilla e intuitiva, incluyendo también un interfaz web.
- El enrutador es modular, flexible y ampliable, permitiendo, en un futuro, migrar a otra tecnología física de conexión a Internet (por ejemplo, fibra óptica) cambiando una de las tarjetas de conexión o añadiendo otra a la infraestructura actual, sin necesidad de desechar ningún equipo.
- El cortafuegos también es modular, permitiendo la contratación futura de distintos productos o módulos software en función de nuevas necesidades.
- Ambos equipos presentan mecanismos de visualización y registro del tráfico que atraviesa sus interfaces, con objeto de solucionar problemas y mejorar el servicio.
- Inclusión de servicios avanzados de seguridad: filtrado con estado, detección y prevención de intrusiones, anti-virus, filtrado URL, etc. que permiten obtener un mayor control en el tráfico de red hacia y desde el exterior (Internet).
- Inclusión de servicios de soporte y mantenimiento: resolución de incidencias de primer nivel, actualización de módulos software y firmware, obtención de parches de seguridad, catálogos de virus y vulnerabilidades, etc.

La nueva infraestructura ya ha sido utilizada para crear un servicio de VPN (Red Privada Virtual) basado en IPSEC que permite a los técnicos del ICANE solucionar incidencias y realizar ciertas tareas de mantenimiento de forma remota. Una vez definido el procedimiento que regulará este servicio, se procederá a ofrecerlo al resto del personal.

### 5.1.3. NUEVAS IMPLANTACIONES DE SISTEMAS

#### 5.1.3.1. Suite de mensajería y colaboración

La necesidad de la implantación de una nueva plataforma de correo electrónico y mensajería surge de los propios objetivos fijados en el PSI para 2009. Un simple análisis de la situación de la plataforma de correo anterior justifica la necesidad del cambio; los componentes de dicha plataforma se mantenían de forma estática, con lo que no reflejaban las mejoras que se han ido incorporando al software con el paso del tiempo. Los usuarios habían elevado quejas debido a la gran cantidad de correo basura (spam) recibido. Ello, unido a la limitación de no poder consultar el correo desde ubicaciones externas al ICANE, justificaba la necesidad de una actualización o mejora de la plataforma.

La plataforma de mensajería "Zimbra Collaboration Suite" fue seleccionada en base a los siguientes criterios:



## ÁREA DE INFORMÁTICA

- Cumplimiento de requisitos tecnológicos y funcionales fijados por el ICANE.
- Garantía de calidad de la plataforma, contando con una entidad sólida respaldándola.
- Soporte técnico directo a través de interfaces web y telefónico.
- Uso amplio de software y estándares abiertos.
- Solución estable con varios años de presencia en el mercado y perspectivas futuras de continuidad.

Su espectro de características es muy amplio, destacando especialmente las que han aportado valor al ICANE en sus procesos de trabajo:

- Acceso web externo y seguro a través de certificado digital a toda la plataforma, incluso desde dispositivos móviles.
- Inclusión de funcionalidades colaborativas avanzadas de correo electrónico, mensajería instantánea local, agenda, planificador de tareas, bloc de notas y maletín de documentos.
- Capacidad para albergar calendarios de eventos compartidos.
- Integración total con clientes Microsoft Outlook (entre otros).

### 5.1.3.2. Arquitectura de integración de datos

El ICANE, dada su naturaleza, recibe datos brutos y elaborados de muy distintas fuentes y en muy diversos formatos. La adecuación, filtrado y limpieza del contenido de ese conjunto de fuentes para su inclusión en los sistemas gestores de bases de datos del ICANE es un proceso largo y tedioso, que requiere de mucha elaboración y especial cuidado.

Se hacía necesaria la implantación de una solución integral de los procesos de extracción, transformación y carga de datos (ETL), permitiendo obtener un repositorio o colección de procesos y trabajos ordenados y correctamente documentados.

Los trabajos realizados en este proyecto se enmarcan dentro del desarrollo del Plan de Sistemas de Información 2009-2011 elaborado por el ICANE. En este contexto, se hallan definidos los siguientes objetivos estratégicos:

- Proposición de soluciones integradas: la concepción global de los sistemas de información garantizará que las soluciones a adoptar en el futuro estén basadas en la integración de las aplicaciones, permitiendo que desde cada Servicio se tenga acceso a una información fiable, actualizada y útil para la actividad que desarrolla.
- Mayor disponibilidad y calidad de la información, mediante el adecuado procesamiento de los datos con distintos orígenes para su posterior mantenimiento de forma integrada y accesible en todas las áreas.

## ÁREA DE INFORMÁTICA

- Mejora de la gestión operativa a través de la automatización de los procesos y procedimientos, ayudando a una mejor realización de los procesos de producción y difusión estadística así como de todas las funciones citadas en los Estatutos del Instituto Cántabro de Estadística.

A su vez, más concretamente y dentro del Plan de Proyectos Anual establecido para 2009, se definieron los siguientes objetivos:

1) Mejorar los procesos de extracción, transformación, carga, calidad e integridad de datos, y a su vez:

- Establecer un sistema automático o semi-automático de carga de datos procedentes de Internet (descargables a través de la herramienta PC-AXIS proporcionada por el INE) y de ficheros sencillos, o directamente de publicaciones recibidas o descargables en formato PDF, de fácil manejo para el personal auxiliar.
- Diseñar una herramienta de edición de estadísticas de síntesis: Anuario Estadístico, Cantabria en Cifras, Fichas Municipales, etc. que permita cargar, en una primera fase, la información recopilada en las bases de datos disponibles. También permitirá al personal completar los datos faltantes.
- Mejorar o reestructurar la herramienta de actualización del sitio web del ICANE, ofreciendo alternativas para subsanar errores de manera que pueda ser utilizada por el personal auxiliar.
- Integrar procesos rutinarios, ofreciendo asistencia para que el personal auxiliar pueda bajar ficheros y procesarlos por software elaborado (Ej.: estadística COMEX).

2) Integración, replicación o unificación del sistema gestor de bases de datos que actualmente da soporte al servidor web y al banco de datos público y el interno, que hace las funciones de almacén de datos.

La solución implantada, basada en el componente de software Pentaho Data Integrator, de la Suite de BI Pentaho, presenta, entre otras, las siguientes características:

- Ofrece soporte en todas las tareas, procesos y trabajos que conforman la extracción, transformación, carga y verificación de los datos recibidos en el ICANE, teniendo como origen fuentes heterogéneas y como destino el almacén de datos.
- Permite la planificación o programación (scheduling) de las tareas desarrolladas, de manera que puedan ser ejecutadas sistemática y automáticamente.
- Permite la creación de un catálogo o repositorio de tareas ETL para su posterior reutilización en distintos motores de datos, de manera que ofrezca una mayor facilidad de mantenimiento y mayor uniformidad en los procesos de carga.
- Ofrece capacidades colaborativas para el equipo de desarrollo.



## ÁREA DE INFORMÁTICA

- Proporciona un amplio espectro de transformaciones disponibles para ser aplicadas a los distintos conjuntos de datos, así como permitir su ampliación de forma independiente de proveedores.
- Presenta sistemas de depuración de procesos y monitorización del rendimiento.
- Está basado en estándares abiertos e interoperables y/o de amplia difusión J2EE.
- Ofrece un interfaz gráfico de usuario sencillo que permite definir los procesos y transformaciones ETL.
- Permite la extracción de datos de todo tipo de fuentes: archivos planos, XML, hojas de cálculo, documentos de texto, bases de datos tanto abiertas como propietarias, RSS, etc.).
- Es escalable, permitiendo un crecimiento controlado del volumen de datos que minimiza el impacto en los procesos ETL.
- Es integrable con otros productos similares o relacionados, como suites o paquetes de minería de datos, análisis OLAP, etc.
- Es modular y extensible, permitiendo el desarrollo de extensiones (plug-ins) a medida en caso de ser necesarios.
- Forma parte de un proyecto de continuidad que asegura el desarrollo de parches, nuevas versiones con nuevas funcionalidades, etc.

Además de realizarse la implantación de la nueva arquitectura, actualmente se encuentra catalogada, documentada y procedimentada prácticamente la totalidad de los procesos de carga y transformación de datos habituales existentes en el ICANE.

### 5.1.3.3. Motor documental

Tanto los procesos de trabajo del ICANE como sus sistemas de información requieren de exhaustiva documentación que permita conservar la experiencia adquirida y que facilite la rotación de personal. Para facilitar estas tareas documentales, se ha puesto en marcha un motor Mediawiki que permite la inclusión de texto a través de un navegador web y su adecuada catalogación, facilitando posteriores búsquedas de información.

La principal utilidad de un motor wiki es la capacidad que proporciona al usuario de crear y mejorar las páginas de forma instantánea, proporcionando una gran libertad de edición a través de una interfaz muy simple. Esto permite la participación de múltiples usuarios en su edición, a diferencia de otros sistemas tradicionales menos flexibles.

Entre las ventajas aportadas al ICANE pueden destacarse las siguientes:

- Redacción de artículos colectivamente (co-autoría) a través de un navegador.
- Facilidad para crear y actualizar nuevas páginas y categorías.

## ÁREA DE INFORMÁTICA

- Posibilidad de revisión y corrección con sellado temporal.
- Identificación y autenticación de usuarios integrada con LDAP corporativo.
- Generación automática de documentos PDF en base a los contenidos introducidos.
- Mejora en la estructuración y catalogación de la documentación asociada a los proyectos así como en la rapidez de las búsquedas de contenido.

### 5.1.4. PROYECTOS DE EXPLOTACIÓN DE DATOS ESTADÍSTICOS

A modo de panorámica, pueden proporcionarse los siguientes datos objetivos relativos a 2009:

- Número estimado de cargas de datos realizadas: 4220 (3395 actualizadas y 59 nuevas).
- Número estimado de transformaciones ETL para series estadísticas: 821.
- Número estimado de transformaciones ETL para microdatos: 281.
- Número estimado de transformaciones ETL para archivos PC-AXIS: 308.

#### 5.1.4.1. Encuesta Marco Input-Output 2007

Esta encuesta, que se realiza por primera vez en 2009, ha requerido el diseño de un nuevo modelo de datos y procesos ETL (extracción, transformación y carga) para la correcta recepción de los cuestionarios enviados por la empresa encargada del trabajo de campo.

Asimismo, se ha diseñado una nueva área temática en el repositorio de la herramienta Oracle Discoverer, para permitir el análisis de la información al Servicio de Estadística.

#### 5.1.4.2. Encuesta Social de Cantabria 2009

Se ha desarrollado de una aplicación a medida para permitir a los encuestados la descarga y posterior envío de los cuestionarios desde/hacia el servidor web del ICANE.

#### 5.1.4.3. Estadística de Comercio Exterior

Se han agregado nuevas clasificaciones de actividades y productos, necesarias para la elaboración de la Contabilidad Regional y el Marco Input-output. En este proceso se han introducido mejoras en el mantenimiento de las tablas de clasificaciones y en los procesos ETL de carga de microdatos mensuales.

Se ha generado la publicación correspondiente al período 2005-2007.

#### 5.1.4.4. Encuesta de Ocupación Hotelera

La Encuesta de Ocupación Hotelera del INE ha sufrido un cambio considerable en la estructura de los ficheros de microdatos motivada por la introducción, a finales de 2008, de un nuevo tipo de cuestionario XML, generado automáticamente por las aplicaciones de los establecimientos hoteleros adheridos al nuevo sistema. La recepción de estos nuevos ficheros ha obligado a



## ÁREA DE INFORMÁTICA

modificar los procesos ETL de carga de microdatos, y la programación destinada a calcular los datos estimados de ocupación hotelera y plazas ofertadas.

Los cálculos de datos estimados, debido a las distintas estratificaciones comarcales empleadas por el INE y el ICANE, requieren el ajuste de los totales por categoría y nacionalidad con los publicados por el INE. Para ello, hasta ahora se seguía un laborioso proceso manual de descarga de información desde la página Web del INE hacia la base de datos Oracle, sin el cual no es posible realizar dicho ajuste. Este proceso se ha automatizado por completo a través de la nueva arquitectura de integración de datos descrita anteriormente.

### **5.1.4.5. Movimiento Natural de Población**

A partir de enero de 2009 han entrado en vigor los nuevos diseños de registro de las estadísticas de Partos y Defunciones, lo que ha obligado a revisar tanto los procesos ETL como los modelos de datos y las capas de usuario final en Oracle Discoverer.

Se ha definido un protocolo de intercambio de ficheros de defunciones entre el INE, la Consejería de Sanidad y el ICANE, que incrementa la seguridad en las transferencias de información.

### **5.1.4.6. Fichas Municipales 2008**

Las Fichas Municipales son una publicación de síntesis, generada a partir de la recopilación de información de ámbito municipal, procedente de fuentes diversas. Se han diseñado procedimientos ETL específicos para los datos municipales.

También se han introducido modificaciones en el diseño de la publicación, debido a la variación de algunas fuentes, así como a la necesidad de actualizar determinadas consultas a la base de datos.

### **5.1.4.7. Directorio de Empresas 2008**

La elaboración del Directorio de Empresas es una operación compleja, que requiere ser revisada anualmente. Recientemente se han comenzado a unificar criterios entre las diferentes Oficinas de Estadística de las CCAA y el INE, fruto de la actividad del grupo de trabajo que coordina el ICANE.

Como resultado de los acuerdos adoptados por dicho grupo de trabajo, previamente a la actualización del Directorio se han realizado las siguientes tareas de depuración:

- Cruce con el Directorio Central de Empresas del INE, para detectar las causas de las diferencias entre ambos.
- Depuración de la codificación de actividades, según la clasificación CNAE 2009.
- Detección de errores en la asignación de epígrafes del IAE.
- Implementación de cambios metodológicos en la programación, en lo referente a los criterios para asignar altas y bajas.

## ÁREA DE INFORMÁTICA

Una vez actualizado el Directorio con los datos correspondientes al año 2008, se ha procedido a su publicación en el Banco de Datos del ICANE.

### 5.2 PROYECTOS EN MARCHA

#### 5.2.1. REESTRUCTURACIÓN DEL SITIO WEB EN UN PORTAL CORPORATIVO

A lo largo del año 2009 se han completado las fases de análisis y diseño del nuevo portal, estando prevista su ejecución durante el primer semestre de 2010.

El sitio web del ICANE representa su principal medio de difusión de información: bancos de datos, publicaciones y noticias estadísticas, trabajos técnicos, etc. Se trata de la referencia estadística de la Comunidad Autónoma de Cantabria, y su existencia responde a la política de difusión establecida según lo dispuesto en los Estatutos del ICANE.

Su evolución en el tiempo es acorde con las necesidades de sus usuarios, tanto internos como externos, así como con las tecnologías vigentes en un momento dado o "estado del arte". De este modo, desde su puesta en marcha se ha abarcado desde la publicación de información estática hasta la gestión de contenidos y la posibilidad de realización de consultas dinámicas a través de desarrollos de aplicaciones web propias en tecnologías Java.

Las razones que motivan la necesidad de iniciar este proyecto se enmarcan en el siguiente contexto:

- Plan Estratégico de Sistemas de Información 2009-2011, elaborado por el ICANE, donde se fija como objetivo prioritario a cumplir durante 2009 la mejora del sitio web actual motivada por las valoraciones realizadas por el personal del ICANE en su conjunto, en términos de lo dispuesto en esta especificación y otros documentos relacionados.
- Plan de Difusión 2009: en este estudio se analiza la principal plataforma de difusión del ICANE, su sitio web, detectándose carencias y puntos mejorables tanto en los servicios ofrecidos como en la organización del contenido.

Se hacen patentes, por tanto, las siguientes necesidades:

- Mejora de la imagen del ICANE como institución de cara a los ciudadanos y entidades tanto públicas como privadas.
- Mejora de la calidad, organización, gestión y acceso de los contenidos ofrecidos.
- Mayor difusión de los servicios ofrecidos por el ICANE.
- Mayor escalabilidad y modularidad, obteniendo un portal ampliable, fácil de mantener y sencillo de administrar.
- Mejora de la relación con ciudadanos, instituciones y empresas, aumentando la afluencia y el acceso a la información ofertada.
- Ampliación de la oferta de servicios a un mayor número de usuarios.



## ÁREA DE INFORMÁTICA

- Ampliación y mejora de la oferta de información y datos personalizados.

Tomando como punto de partida lo expuesto, se plantean las siguientes metas:

- Estar en disposición de obtener una certificación AENOR de conformidad con los requisitos de la Norma de referencia (UNE 139803) de nivel AA, permitiendo la posibilidad de conseguir, en un futuro, la marca N de accesibilidad TIC de AENOR.
- Estar en disposición de integrar el portal con componentes de Suites de Inteligencia de Negocio (Business Intelligence) disponibles en el mercado.
- Mejorar la oferta, calidad y gestión de contenidos, traducándose en :
  - Un aumento del número medio de accesos al nuevo portal.
  - Mantenimiento de la constancia o frecuencia de dichos accesos.
  - Reducción del tiempo empleado en la inclusión y gestión de los contenidos por parte de los técnicos del ICANE.

### 5.2.2. REESTRUCTURACIÓN DEL BANCO DE DATOS

A lo largo del año 2009 se han completado la fase de análisis del nuevo sistema banco de datos, estando prevista su ejecución durante el primer semestre de 2010.

El banco de datos del ICANE conforma uno de los servicios de mayor relevancia publicados en su sitio web. Se ideó con el objeto de ampliar la oferta de información estadística al público interesado, poniendo a su disposición un sistema de consultas interactivas. Su funcionalidad reside en ofrecer a sus usuarios la posibilidad de realizar dichas consultas "a medida", simplificando el acceso y comprensión de determinada información que, por su complejidad técnica, resulta difícil de tratar por el público no especializado.

La evolución del banco de datos del ICANE en el tiempo ha quedado relativamente estancada en su primera versión, desarrollada principalmente sobre tecnología Java (archivos JSP planos). Su interfaz no es lo que un usuario de aplicaciones web espera hoy en día, pudiendo ser claramente mejorable. Del mismo modo, su gestión y mantenimiento son costosos y su potencia de consulta ha quedado limitada tras la consolidación de nuevas tecnologías de análisis exhaustivo de datos.

Las razones que motivan la necesidad de iniciar éste proyecto se enmarcan en el siguiente contexto:

- Plan Estratégico de Sistemas de Información 2009-2011, elaborado por el ICANE, donde se fija como objetivo prioritario a cumplir durante 2009 la mejora del banco de datos actual motivada por las valoraciones realizadas por el personal del ICANE en su conjunto, en términos de lo dispuesto en ésta especificación y otros documentos relacionados.
- Plan de Difusión 2009: en este estudio se analiza la principal plataforma de difusión del ICANE, su sitio web, y dentro de éste, su banco de datos.

## ÁREA DE INFORMÁTICA

Se han detectado carencias y puntos mejorables en cuanto a las capacidades de análisis estadístico, los cuales dificultan su difusión a un mayor número de usuarios.

Se hacen patentes, por tanto, las siguientes necesidades:

- Mejora del interfaz de usuario.
- Mejora del sistema de consulta, ofreciendo mayor potencia en la selección, cruce y representación de variables.
- Mejora en los sistemas de búsqueda dentro del banco de datos.
- Mayor escalabilidad, facilidad de mantenimiento y de gestión.
- Adecuación del modelo de datos al nuevo sistema de consulta.
- Ampliación de la oferta de información a un mayor número de usuarios.

Mediante la ejecución de este proyecto se pretende conseguir las siguientes metas:

- Un modelo de datos relacional normalizado, para los Bancos de Datos de Series Temporales, Series Históricas y Series Municipal.
- Un modelo de metadatos extendido y normalizado, así como de un sistema de gestión y mantenimiento de los mismos.
- Un sistema de análisis multidimensional explotado mediante tecnología OLAP.
- Una reducción de la complejidad de utilización del sistema en su conjunto, traducido en un aumento del número total de usuarios.
- Una reducción del número de peticiones de datos a medida, dotando al usuario del sitio web de mayor capacidad para ejecutarlas por sí mismo.

### 5.2.3. AUTOMATIZACIÓN DEL SISTEMA DE PUBLICACIONES

Como parte de su política de difusión, el ICANE genera del orden de aproximadamente 2.500 páginas anuales de publicaciones estadísticas. El proceso de edición es complejo, principalmente debido a la gran cantidad de datos tabulados y cruzados que han de incluirse en las publicaciones, en mayor o menor medida.

El proyecto de automatización se encuentra actualmente en su fase de análisis, estando programado el inicio de ejecución en 2010. A lo largo del presente año se ha realizado una especificación de requisitos del sistema y un estudio de viabilidad de varias posibles soluciones, así como una propuesta final. El sistema debe mejorar la eficacia de las actividades realizadas en el proceso de producción de contenido estadístico, permitiendo una reducción del tiempo empleado en su elaboración, y por tanto, del coste de producción de dicho contenido. Se espera que proporcione los siguientes beneficios:

- Liberar al personal estadístico de la carga de maquetación de contenidos.



## ÁREA DE INFORMÁTICA

- Aumentar la independencia del personal estadístico con respecto del personal informático para la modificación de publicaciones.
- Mejorar la calidad de las publicaciones y la automatización del proceso con respecto a los sistemas actuales.
- Reducción del tiempo de realización de publicaciones y coste de producción de contenido.
- Aumento de la flexibilidad para tratar posibles cambios en las publicaciones.

Asimismo, se espera alcanzar las siguientes metas:

- Incrementar al máximo el grado de automatización de toda la producción de contenido para publicar, especialmente aquellos procesos realizados en la actualidad de forma "semi-automática".
- Migrar aquellos procesos ya automatizados mediante otros sistemas que sean susceptibles de ser mejorados a través del nuevo sistema.

