



Comunicación

148

MERLIN: SISTEMA DE ANALISIS DE LA INFORMACIÓN DEL SPEE-INEM. ESTRUCTURA Y POSIBILIDADES DE USO PARA EL ESTUDIO DEL MERCADO LABORAL

Antonio Hernando Sánchez

Jefe de Área de Sistemas de Análisis de la Información
Subdirección General de Informática y Estadística
Servicio Público de Empleo Estatal-INEM

Palabras clave

Sistemas de análisis de la información, MERLIN, mercado laboral, INEM-SPEE, Servicio Público de Empleo Estatal, Sistema Informático de los Servicios Públicos de Empleo, SISPE.

Resumen de su Comunicación

En esta comunicación se presenta un breve recorrido por el denominado MERCado Laboral del INEM (MERLIN), Sistema de Análisis de la Información (SAI) del SPEE-INEM, que recoge desde la solución técnica adoptada actualmente, pasando por el tipo y volumen de datos que se tienen, hasta las posibilidades de su uso para el estudio del mercado laboral.

MERLIN: SISTEMA DE ANÁLISIS DE LA INFORMACIÓN DEL SPEE-INEM. ESTRUCTURA Y POSIBILIDADES DE USO PARA EL ESTUDIO DEL MERCADO LABORAL

0. Introducción

Los sistemas de análisis de la información son algo más que un conjunto de datos históricos almacenados de forma adecuada para su consulta. Con ellos el analista de la información o técnico especialista de la unidad, sin necesidad de colaboración de terceros, mediante el uso de herramientas amigables y de fácil uso, consigue autoabastecerse de los datos necesarios para sus procesos de análisis con el fin de evaluar acciones pasadas o definir nuevas estrategias.

En esta comunicación se presenta un breve recorrido por el denominado MERCado Laboral del INEM (MERLIN), Sistema de Análisis de la Información (SAI) del SPEE-INEM, que recoge desde la solución técnica adoptada actualmente, pasando por el tipo y volumen de datos que se tienen, hasta las posibilidades de su uso para el estudio del mercado laboral.

Los sistemas de análisis de la información con datos procedentes de los registros administrativos, estén cobrando cada vez más relevancia para estudios estadísticos complementarios de las encuestas e incluso pueden ser sustitutorios de aquellas en los ámbitos geográficos pequeños, que las muestras no son fiables o no existen. El hecho de tener grandes bases de datos longitudinales que recogen todo el universo al que se refieren, junto con herramientas analíticas cada vez más eficientes y de más fácil uso, abren unas perspectivas ilimitadas para el tratamiento automático de la información por la facilidad de tratamiento de los datos, unido al bajo coste de obtención de los mismos.

1. MERLIN: El sistema de análisis de la información del SPEE-INEM.

1.1. El Sistema de Información de los Servicios Públicos de Empleo

El SAI del SPEE-INEM está enmarcado dentro del Sistema de Información de los Servicios Públicos de Empleo (SISPE).

Para garantizar que los distintos servicios públicos de empleo puedan seguir desempeñando las funciones que les han sido encomendadas, el Sistema Nacional de Empleo, que está integrado por el SPEE-INEM y los respectivos servicios de públicos de empleo de las comunidades autónomas, asigna al proyecto SISPE la función de integrar toda la información disponible sobre la intermediación laboral y las políticas activas de empleo. Mediante este sistema se consigue además que todos los sistemas parciales de las comunidades estén integrados en uno solo y que los trabajadores y empleadores perciban un único mercado laboral de ámbito nacional.

El SISPE está construido de forma totalmente abierta habiéndose definido, por un lado, un conjunto de datos comunes con las mismas claves, y por otro, los procedimientos de gestión comunes consensuados entre todos los servicios públicos de empleo. Se consiguen así una serie de entidades y procesos comunes que cada sistema autonómico puede implementar físicamente como considere conveniente, además de poder añadir datos y procesos propios de su comunidad, siendo transparente para el intercambio de información entre los distintos sistemas.

Con este sistema cada solicitante de un servicio: empresario que realiza una oferta o trabajador que pide un empleo, puede hacer sus gestiones desde cualquier oficina de cualquier servicio público de empleo, independientemente del lugar donde se encuentre el centro de trabajo.

Cómo principales características del SISPE se pueden citar:

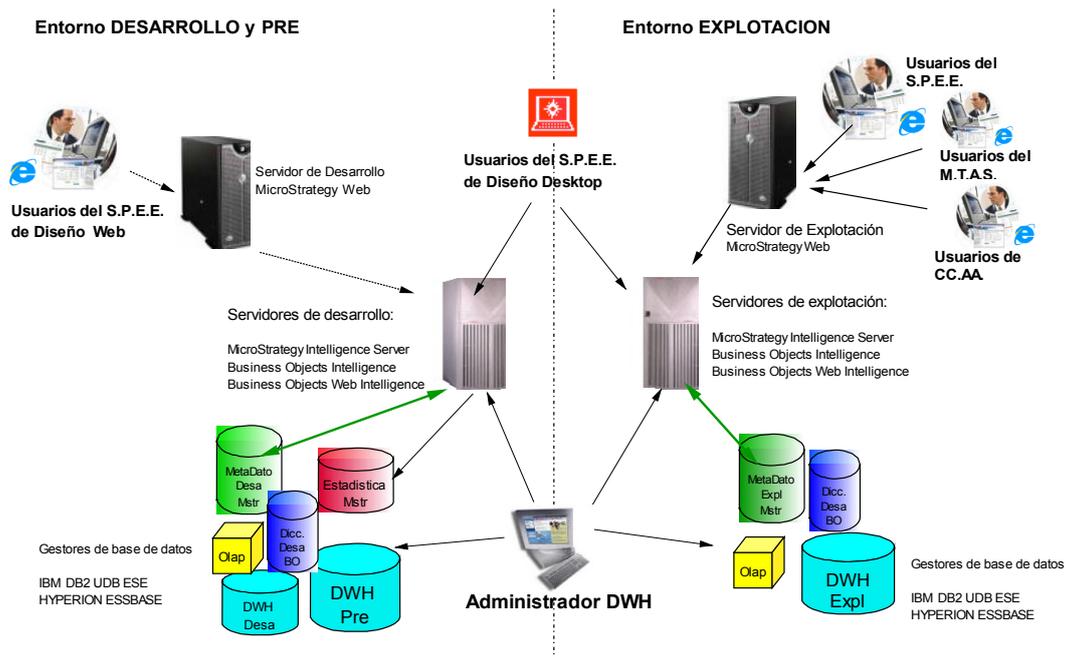
- Todos los sistemas están interconectados a través del sistema informático del SPEE.
- En la base de datos central existe una copia actualizada en tiempo real de todos los datos comunes de los sistemas individuales de cada comunidad.
- Cualquier dato puede ser consultado y actualizado desde cualquier sistema comunitario.
- El acceso de la información entre comunidades se hace a través del sistema del SPEE.

Esta definición de los datos y su tratamiento: datos comunes y con los mismos códigos y única definición de procesos de gestión, garantizan la integridad de cada dato independientemente de su lugar de procedencia, siendo este hecho la base fundamental de la importancia del SAI del SPEE-INEM para realizar estudios sobre el mercado laboral: haciendo que éstos sean totalmente homologables cuando se refieren a cualquier ámbito geográfico y el resultado a nivel nacional sea la suma de los resultados respectivos de las comunidades.

Conviene también destacar que durante la definición de requerimientos del nuevo sistema hubo siempre una estrecha colaboración entre las unidades gestoras y las dedicadas a estudios analíticos para que aquellas, durante la gestión, hicieran la toma de datos necesarios para los estudios sobre el mercado laboral.

1.2. Arquitectura física

El sistema MERLIN es un SAI clásico, separado del sistema operacional y su arquitectura, cuadro 1, está formada por un servidor SUN FIRE 15K con 40 CPU's UltraSPARC-III y 104 GB de RAM, (32 CPU's y 84 GB de RAM del que son de uso exclusivo para el SAI) y sistema operativo SUN Solares 9.



Cuadro 1: Arquitectura física de MERLIN

La partición dedicada al MERLIN consta de dos dominios, el de explotación, con 20 CPU´s y 48 GB de RAM, y el de desarrollo y preexplotación, con 12 CPU´s y 36 GB de RAM.

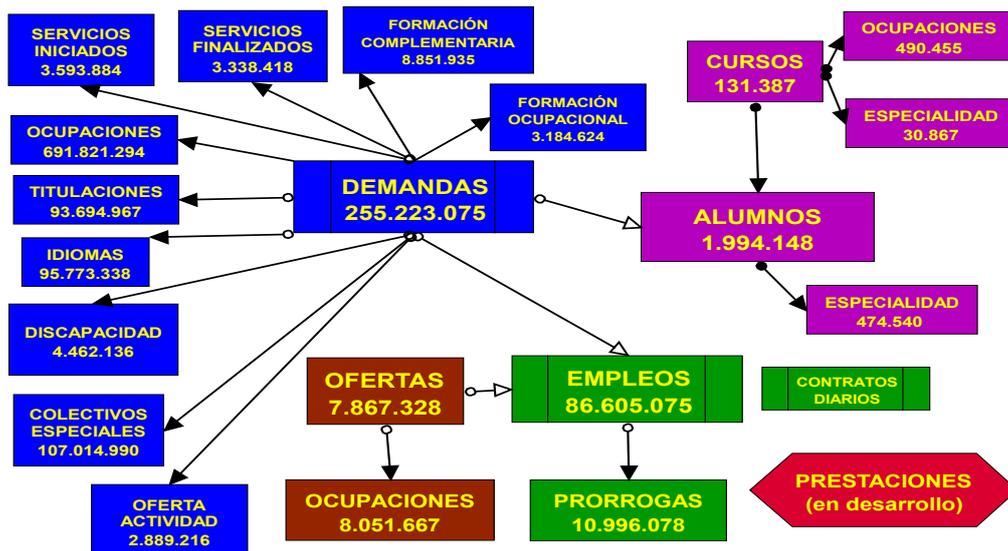
Los datos se encuentran almacenados, al máximo nivel de detalle, con el gestor relacional IBM DB2 UDB ESE v8.1 FP8a (arquitectura de 64 bits) y bajo el gestor multidimensional HYPERION ESSBASE datos agregados según las variables de interés estadístico. Los datos que se cargan son procedentes del sistema de gestión y por tanto tienen que sufrir transformación para su adaptación a las bases de datos de consulta. Las cargas incrementales se hacen una vez al mes.

Se tienen dos tipos de herramientas de acceso para usuario final: MICROSTRATEGY, para 250 usuarios analistas con capacidad de ejecutar cualquier informe en la base de datos relacional, y BUSINESS OBJECTS con 500 usuarios para consultas a la base de datos relacional o multidimensional. Ambas herramientas están actualmente instaladas con accesos web a través de servidores WINDOWS¹.

1.3. Tipo de datos

Como se ha dicho los datos son procedentes de los servicios públicos de empleo y se almacenan los relativos a los últimos cinco años.

¹ Se ha comenzado el proceso de migración de los servidores de las herramientas de usuario desde WINDOWS al sistema UNIX.



Cuadro 2: Modelo de datos

Como se puede apreciar en el Cuadro 2, que también contiene el número de registros de cada tabla, se almacena información sobre las principales entidades del INEM: Demandas, Ofertas, Empleo, Formación.

La información se almacena a nivel de detalle, con el fin de poder realizar análisis que relacionen todas las entidades.

Para mantener la calidad en los datos almacenados, antes de cargar la información, esta sufre un proceso de depuración y adecuación con el fin de que ésta sea coherente (debe cumplir integridad referencial, todos los códigos estarán registrados) y por tanto los estudios a realizar sobre ella también lo serán.

El modelo de datos está basado en un modelo multidimensional (copo de nieve), que permite navegar por la información a un mayor o menor nivel de detalle, o a otros puntos de vista del negocio.

2. Entornos de consulta de MERLIN

Hay tres entornos de acceso con un único punto de entrada para los usuarios.

2.1. Estadísticas

Las estadísticas se encuentran almacenadas en el gestor multidimensional y se refieren a todas las áreas de actuación del SPE, el usuario puede obtener cualquier informe de la serie estadística sobre el mercado laboral referido a cualquier mes de los últimos años, restringido al ámbito geográfico que desee y a través de múltiples filtros para las otras variables estadísticas de forma prácticamente inmediata y en cualquiera de los formatos ofimáticos principales. Hay varios tipos de acceso: mediante un navegador de internet usando la herramienta BUSINESS OBJECTS o mediante la red local desde hojas de cálculo.



Cuadro 3: Pantalla de acceso a MERLIN

2.2. Sistema de análisis de la información

Entorno de análisis para usuario analista, principalmente de los observatorios de empleo, que permite realizar libremente consultas complejas, dinámicas y no predefinidas con cualquier nivel de detalle en el gestor relacional [Datos históricos de cinco años]. Se usa la herramienta MICROSTRATEGY para este entorno. Estos usuarios pueden diseñar sus propios informes y compartirlos con los usuarios de su grupo. Los proyectos destinados a este grupo están orientados más a la estrategia que a la operativa de gestión.

2.3. Informes del mercado laboral

Aplicación definida para que los técnicos de empleo, basándose en un conjunto de informes predefinidos, puedan hacer un estudio del mercado laboral referido al ámbito geográfico de su actuación orientado a la gestión.

3. Posibilidades de estudios para el análisis del mercado de trabajo.

Como se ha visto los datos básicos de MERLIN proceden de registros administrativos, y tienen las virtudes y los inconvenientes propios de este tipo de fuentes. Su mayor ventaja es la enorme cantidad de observaciones, que hace que sean significativos a cualquier nivel de desagregación. Por ello son especialmente útiles para estudios de zonas geográficas pequeñas, donde las encuestas de ámbito nacional no dan resultados fiables o no existen. Otra de sus ventajas es la rapidez con que se computan: los datos básicos se hacen públicos el segundo día del mes siguiente. Las bases de datos del SPEE contienen

información sobre una misma persona en períodos diferentes, lo que abre posibilidades de utilización todavía no suficientemente exploradas.

Otra gran ventaja de los registros de los servicios públicos de empleo es que su información no está basada solo en lo que declaran los inscritos, es acreditada documentalmente en el momento de ser solicitada (datos de formación de los demandantes, datos sobre reconocimiento de prestaciones,...), también es depurada a través de cruces con otras unidades externas o de otras áreas (estudiantes, pensionistas, ocupados,...), y por los gestores en los contactos con los demandantes. Al existir una única base de datos central con todos los demandantes, que pueden consultar todas las comunidades, se impide que existan duplicidades de las mismas demandas en distintas comunidades. Antes de inscribir a un demandante se verifica que no está inscrito en ninguna comunidad.

También hay que destacar el hecho, posiblemente único en la Administración en materias que tiene la gestión descentralizada, que todos los registros públicos de empleo de las distintas comunidades usen las mismas claves e idénticos procesos de tratamiento, lo que hace que los resultados estadísticos sean totalmente equiparables en todas las comunidades y el dato nacional sea la suma de todos los comunitarios.

Se tiene una perspectiva longitudinal, que permite hacer estudios dinámicos a lo largo del tiempo, incluso referidos a un mismo individuo. Todos los hechos registrados están enlazados a través de los identificadores de las personas/empresas a las que se refieren.

Entre los estudios que se realizan en el sistema MERLIN, además de los estudios de tipo estructural, cabe destacar:

- Estudios del empleo a través de los contratos registrados:

A pesar de que el registro de contratos sólo se recoge el momento de entrada y se tiene información posterior y la duración sólo se conoce para un porcentaje de los contratos temporales. Tiene interés, no sólo en el análisis transversal de los contratos, si no también en la posibilidad de hacer el seguimiento a lo largo del tiempo. Los principales estudios están referidos a la inserción de los trabajadores después de un episodio de desempleo o la inserción laboral de los alumnos de formación ocupacional. Estudios del segundo tipo pueden ser la recurrencia de los trabajadores en la temporalidad hasta su inserción en la contratación indefinida o la reincidencia de contratación por las mismas empresas, movilidad geográfica de los trabajadores,...

- Estudios de desempleo a través del registro de demandantes de empleo:

La tabla de hechos de demandantes de empleo es un fichero vivo del que se guarda una imagen de la situación de los demandantes que abandonan el registro durante el mes y de los que permanecen el último día del mes y por tanto pasan como demandantes al mes siguiente. Esto permite analizar todas las entradas y salidas del registro así como el stock de los que permanecen en el cambio de mes, además de poder hacer un seguimiento de los demandantes de empleo a lo largo del tiempo, relacionándolos con todos los hechos que haya tenido referidos a contrataciones, cursos de formación recibidos, servicios orientados a la reinserción en el mercado laboral, prestaciones por desempleo percibidas, ofertas de empleo con las que ha sido relacionado,...

- Otro tipo de estudios:

Evaluación de las políticas activas de empleo, descubrimientos de nuevos yacimientos de empleo, análisis del alcance de nuevas estrategias para la inserción laboral,...

4. Perspectivas de futuro

Como se ha visto la posibilidad del tratamiento masivo de la información abre nuevos canales de explotación estadística de los registros administrativos que podría complementar e incluso sustituir en algunos casos a las operaciones de campo de las estadísticas, por representar mayor fiabilidad y menor coste. Por ello es necesario una colaboración entre las unidades de gestión y las dedicadas a los estudios estadísticos o analíticos. Así mismo, no se deben escatimar esfuerzos y recursos a la explotación de los registros administrativos adecuándolos con los datos que sean necesarios para los fines analíticos, aunque sean de poca o nula importancia para la gestión del fin al que está destinado el registro.

También se necesita una colaboración entre las unidades gestoras que permita cruzar registros administrativos de distinta procedencia para la obtención de estudios e incluso sería deseable interconectar sistemas de análisis de la información que permitan estudios mas complejos que los que se pueden hacer actualmente referidos a sistemas independientes de cada unidad. Por ejemplo, para el estudio del mercado de trabajo: empleo, desempleo e ingresos salariales; sería de interés cruzar datos procedentes de la Seguridad Social, para establecer datos sobre el empleo y los ingresos salariales, con datos de los servicios públicos de empleo: contratos, demandantes de empleo y perceptores de prestaciones por desempleo y con datos de los registros tributarios que permitirían calcular los ingresos salariales.