## Cómo garantizar la calidad y limpieza de datos

Una tarea imprescindible para asegurarse de que los datos son apropiados para su uso empresarial es el control de calidad y la limpieza de los datos. Esta tarea está normalmente delegada entre el departamento de TI y las áreas funcionales que son responsables del mantenimiento de cada base de datos.

El departamento de TI debe responsabilizarse de la integridad y la calidad de los orígenes de los datos.

Por otro lado, las áreas funcionales disponen del conocimiento necesario sobre sus datos para limpiar datos, que se puede resumir en buscar coincidencias (duplicidades) y mantener el formato de los datos.

Hay que tener siempre en cuenta que estas tareas afectan tanto a las bases de datos de negocio como a las tablas maestras. En ambos casos se han definido qué campos son obligatorios, o cuáles no pueden repetirse el valor (por ejemplo el NIF/VAT de una empresa), lo que ayuda de forma directa a evitar datos corruptos, por ejemplo, empresas duplicadas.

Por lo tanto, se han definido unas tareas planificadas (TP3 y TP6) para la limpieza de datos de forma periódica.

Además, al CRM se le requerirá una serie de herramientas que ayuden al control de las duplicidades; por ejemplo, en la BD2 Contacts deberán existir reglas que revisen ciertos datos que ayudan a la identificación de duplicidades de los contactos. Los proveedores deberán proporcionar esas reglas y adaptarlas a las bases de datos de IPT-FV.

## Funciones de las bases de datos

Para la explotación y mantenimiento de las bases de datos se han definido las siguientes funciones:

* Exportación de cualquier base de datos o tabla maestra a los siguientes formatos de fichero:
  + CSV ISO-8859-1 (estándar para su uso en herramientas ofimáticas)
  + CSV UTF-8 (estándar para su uso en medios de correo electrónico e internet)
  + SQL
  + XML
  + MS Excel, en formato de Microsoft DOCX
  + PDF
* Importación de ficheros. Se requerirá de una herramienta de importación que permita, al importar un fichero, seleccionar qué campos importar y qué campos de destino se les asigna. La importación se podrá realizar desde los siguientes formatos de fichero:
  + CSV, pudiendo seleccionar separador de campos
  + SQL
  + XML
  + MS Excel, en formato de Microsoft DOCX