Digital Signature Calculator (DSC) 3.1

Manual del Usuario

LabCert PUCP
Av. Universitaria 1800
Lima - Perú

Enero 2006

Tabla de Contenidos

1 Introducción

Este manual describe el uso y las funciones del programa Digital Signature Calculator (DSC), versión 3.1.

DSC es un programa que calcula la huella digital (*signature*) de archivos almacenados en una computadora. A través de la huella digital de un archivo se puede conocer el estado de su integridad, comparando el resultado obtenido con otro previamente conocido. Si ambos resultados concuerdan podemos decir que el archivo analizado no ha sufrido alteraciones o que es una copia idéntica del archivo maestro utilizado para establecer el patrón.

La huella digital es un número que suele ser expresado como cadena de caracteres hexadecimales, calculado utilizando toda la información binaria contenida en el archivo (que a su vez puede ser interpretada como una secuencia de números, sin importar el contenido o la naturaleza de la información guardada en el archivo) y una fórmula matemática que utilice toda la información del archivo. La fórmula matemática puede ser elegida, de manera que tendremos muchos tipos de huellas digitales. Cada uno de estos tipos de huellas digitales ofrecerá diferentes grados de efectividad en su misión de ayudarnos a determinar la validez de un archivo. A esta fórmula matemática se le llama también Algoritmo, por cuanto va a ser implementada en un programa de computadora.

1.1 Descripción de los algoritmos implementados en DSC

Entre los algoritmos más utilizados y presentes en DSC para calcular huellas digitales tenemos:

- **Checksum**. Estrictamente hablando, este algoritmo matemático no es considerado como un algoritmo útil para el cálculo de huellas digitales. Consiste en la suma simple de cada byte que compone el archivo, truncado en un número específico de bits (por ejemplo, 16 o 32). El resultado final de toda la suma es el número que se considera como huella digital. Es muy fácil de implementar y muy rápido debido a que una de las instrucciones más elementales de cualquier computadora es la suma. Sin embargo, no es seguro porque una de las propiedades básicas de la suma es la no

importancia del orden de los sumandos en el resultado final. Un byte errado o alterado puede ser compensado en otro lugar y la suma no se altera. Es utilizado en situaciones en donde la velocidad de cálculo es lo más importante y en donde se pueden tolerar fallas en archivos alterados que son reportados como válidos por el algoritmo.

- **CRC-32**. Significa Checksum Redundante Cíclico de 32 bits. Consiste en operaciones realizadas bit a bit a grupos de datos provenientes del archivo bajo análisis. La huella se va calculando considerando el residuo de una división binaria de la cadena de datos con una cadena predeterminada llamada "polinomio". Este algoritmo es bastante rápido porque se implementa con funciones básicas de cómputo como la suma y las operaciones bit a bit. Tiene un gran uso en los protocolos de telecomunicaciones al validar tramas relativamente cortas de datos.

- **Validator-32**. Es una variante del CRC-32 ofrecida por la empresa Dataman en sus equipos de cálculo de huellas digitales de memorias EPROM.

- **SHA-1**. Es un algoritmo de cálculo de huellas digitales creado para servir como un mecanismo confiable de validación de datos. Fue creado por NIST y presentado como el Estándar FIPS-180. Consiste en una huella digital compleja de 140 bits, Calculada mediante un algoritmo que es de conocimiento público. Es el medio más efectivo

1.2 Características presentes en DSC v3.01

- Versión para Windows 9x/Me/2000/XP.
- Disponible en forma gratuita.
- Calcula Checksum, CRC-32, Validator-32 y SHA-1.
- Genera reportes en archivo de texto, en texto simple o espaciado por tabulaciones.
- Consulta de resultados obtenidos contra una base de datos.
- Inserción de los resultados obtenidos en una base de datos.

1.3 Aviso importante sobre el acceso a bases de datos

Lea cuidadosamente las instrucciones referentes al acceso a bases de datos. El uso indebido de este programa podría causar daños a la base de datos utilizada, generar registros innecesarios o duplicar los mismos. Debido a que los procesos referidos a las bases de datos se realizan en

bloques, es probable que alguna información sea agregada cuando ya existía en la base de datos.

Para utilizar las características de acceso a bases de datos es necesario contar con los permisos necesarios de lectura (para comparación de huellas digitales) y escritura (para añadir registros a la base de datos). Esos permisos (usuarios y claves de acceso) deben ser entregados por el administrador de la base de datos.

La configuración del programa para realizar consultas e inserción de campos en la base de datos requiere información acerca del diseño de la misma. Consulte al administrador de su base de datos acerca de los nombres de tablas, campos y del diseño en general antes de intentar utilizar estas características de DSC.

1.4 Términos utilizados

- **Huella digital**. Número que suele ser expresado como cadena de caracteres hexadecimales que se interpreta como el resumen o cantidad representativa de todo un archivo de datos.
- **Checksum**. Algoritmo que consiste en la suma de los bytes de un archivo.
- **CRC-32**. Algoritmo de verificación utilizado en telecomunicaciones.
- **Validator-32**. Variante del algoritmo CRC-32 presentado por la empresa Dataman en sus equipos verificadores de memorias.
- **SHA-1**. Algoritmo para huellas digitales creado por NIST.
- **Base de datos**. Colección organizada de datos. Las más utilizadas son las bases de datos relacionales, formadas por tablas definidas por campos. La información contenida en los campos para un mismo elemento definen un registro. Se puede visualizar a los campos como las columnas de una tabla y a los registros como las filas.
- **Permiso de acceso**. Elementos secretos (usuario-contraseña) que permiten el acceso a cierta información.
- **Registros de datos**. Grupo de campos de datos que componen la unidad independiente más pequeña y completa de datos.
- **Campos de datos**. Cada una de las piezas de información. Cada campo tiene una estructura y significado independiente.

2 Cálculo de huellas digitales y opciones básicas

El cálculo de huellas digitales con DSC es muy sencillo. La interfase de usuario contiene todas las opciones disponibles en forma de botones:

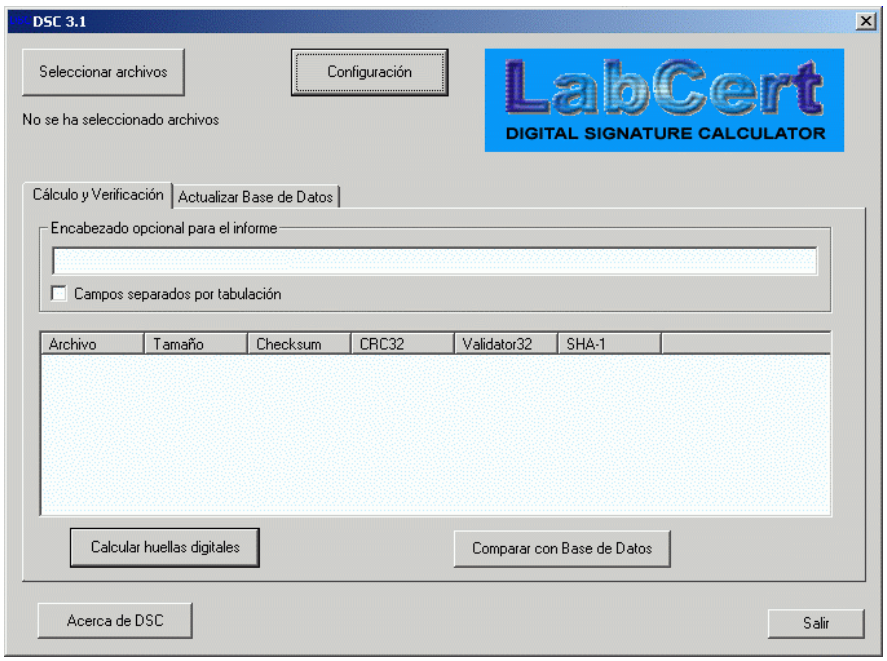


Figura 1. Ventana principal de DSC

2.1 Cálculo de huellas digitales

Para el cálculo simple de huellas digitales debe seguir los siguientes pasos:

- Seleccione la solapa “Cálculo y Verificación”.
- En el campo “Encabezado opcional para el informe” puede ser llenado con cualquier texto. El reporte final que se genera incluirá el contenido de esta línea.
- La opción “Campos separados por tabulación” afecta la manera en que se generará el reporte final en archivo de texto. Cuando esta opción no está marcada se genera un reporte simple en el que se presentan los resultados en varias líneas. Cuando la opción está

marcada, se genera una sola línea para cada archivo analizado, con los resultados de cada huella digital separados por símbolos de tabulación. Este último formato es útil cuando se quiere importar estos datos a alguna aplicación, como por ejemplo, hojas de cálculo.

- Presione el botón “Seleccionar archivos”, una ventana de diálogo estándar del sistema aparecerá para seleccionar los archivos a analizar. Puede seleccionar más de un archivo a la vez, utilizando las teclas SHIFT o CTRL mientras realiza la selección.

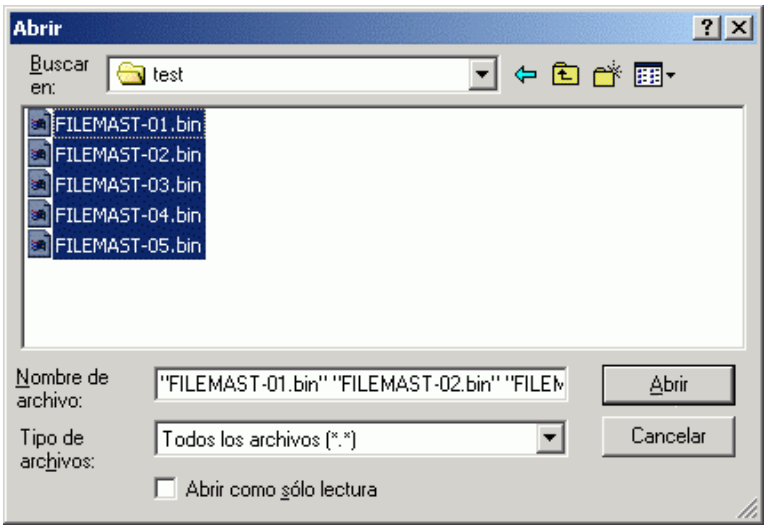


Figura 2. Selección de archivos

- En este punto aparecerá un texto indicando el número de archivos que se han seleccionado y la carpeta. Basta con presionar el botón “Calcular huellas digitales” para empezar con los cálculos.

El tiempo de procesamiento depende del tamaño de los archivos seleccionados y de la velocidad de la computadora que ejecuta la aplicación.

- Cuando la aplicación termina el cálculo de las huellas digitales se muestran los resultados en pantalla de la siguiente manera:

  - Nombre del archivo seleccionado.
  - Tamaño en bytes.
  - Checksum.
  - CRC-32.
  - Validator-32.
  - SHA-1

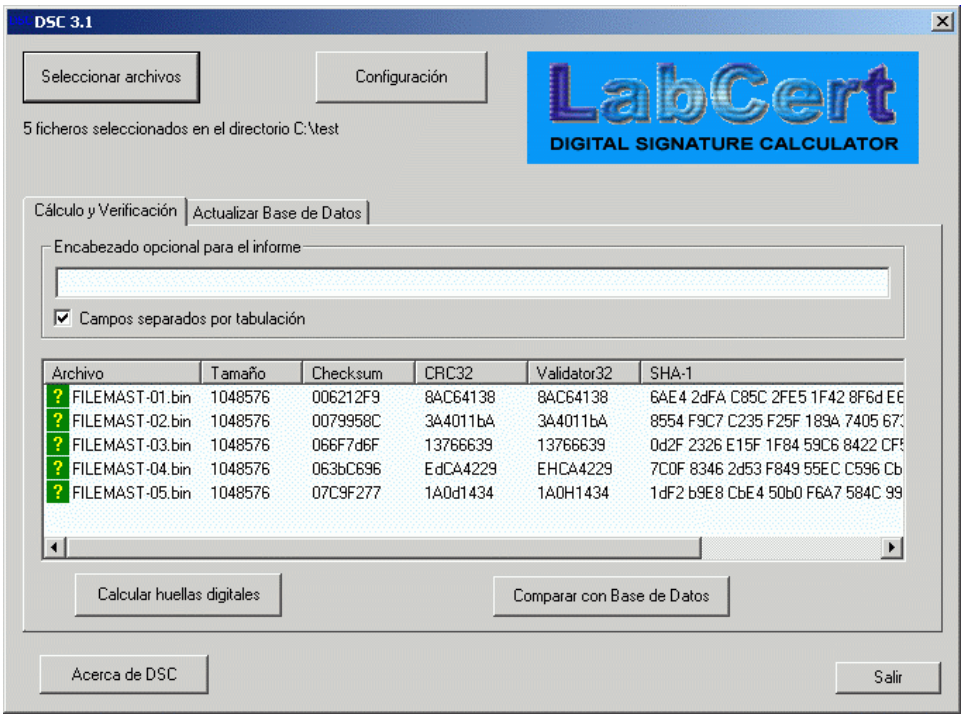
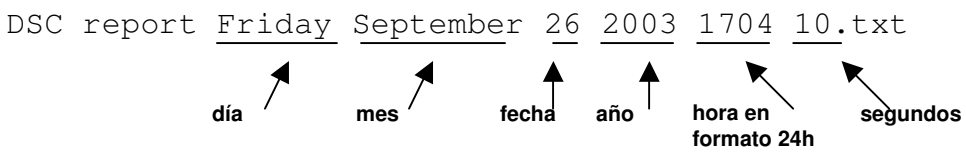


Figura 3. Finalización del cálculo de huellas digitales

El símbolo de interrogación acompaña a cada resultado obtenido porque hasta este punto, no se ha comparado la huella digital obtenida con ningún otro tipo de datos. Pese a esto, el resultado mostrado corresponde a las huellas digitales de los archivos seleccionados.

2.2 Verificando el archivo de reporte

En la carpeta donde se encuentran los archivos seleccionados se generará un reporte en la forma de un archivo de texto. El nombre generado para el archivo de texto tendrá la siguiente estructura:



Esta estructura garantiza que el archivo no se sobrescribirá si se repite el cálculo sobre archivos contenidos en la misma carpeta.

El contenido del archivo de reporte dependerá del estado de la casilla "Campos separados por tabulación" al momento de efectuarse los cálculos. Si la casilla no estaba marcada se genera un reporte simple similar al que se muestra:

```
Archivo:    FILEMAST-01.bin
Tamaño:     1048576
Checksum:   006212F9
CRC32:      8AC64138
Validator32:8AC64138
SHA-1:      6AE4 2dFA C85C 2FE5 1F42 8F6d E652 08CE 34d2 1dCC

Archivo:    FILEMAST-02.bin
Tamaño:     1048576
Checksum:   0079958C
CRC32:      3A4011bA
Validator32:3A4011bA
SHA-1:      8554 F9C7 C235 F25F 189A 7405 673d 13b7 3C23 331b

Archivo:    FILEMAST-03.bin
Tamaño:     1048576
Checksum:   066F7d6F
CRC32:      13766639
Validator32:13766639
SHA-1:      0d2F 2326 E15F 1F84 59C6 8422 CF5A 6d00 AEbA F2Cd
```

Si la casilla "Campos separados por tabulación" estaba marcada, se generará un informe en el que los campos se encuentran separados por signos de tabulación. Este formato es útil cuando se quiere importar estos datos a alguna aplicación, como por ejemplo, hojas de cálculo:

```
Archivo     Tamaño      Checksum    CRC32 Validator32 SHA-1
FILEMAST-01.bin   1048576     006212F9    8AC64138    8AC64138
      6AE4 2dFA C85C 2FE5 1F42 8F6d E652 08CE 34d2 1dCC
FILEMAST-02.bin   1048576     0079958C    3A4011bA    3A4011bA
      8554 F9C7 C235 F25F 189A 7405 673d 13b7 3C23 331b
FILEMAST-03.bin   1048576     066F7d6F    13766639    13766639
      0d2F 2326 E15F 1F84 59C6 8422 CF5A 6d00 AEbA F2Cd
FILEMAST-04.bin   1048576     063bC696    EdCA4229    EHCA4229
      7C0F 8346 2d53 F849 55EC C596 Cb1A FF8F 032C 837F
FILEMAST-05.bin   1048576     07C9F277    1A0d1434    1A0H1434
      1dF2 b9E8 CbE4 50b0 F6A7 584C 998C dA71 14d5 F702
```

Si la carpeta en donde se encuentran los archivos seleccionados tiene establecida la propiedad "Sólo lectura" (como es el caso de las carpetas almacenadas en CD-ROMs) no se generará el archivo de reporte.

2.3 Configurando el idioma de operación

El botón “Configuración” presenta la ventana general de configuraciones, en donde se pueden modificar algunos parámetros de operación de la aplicación.

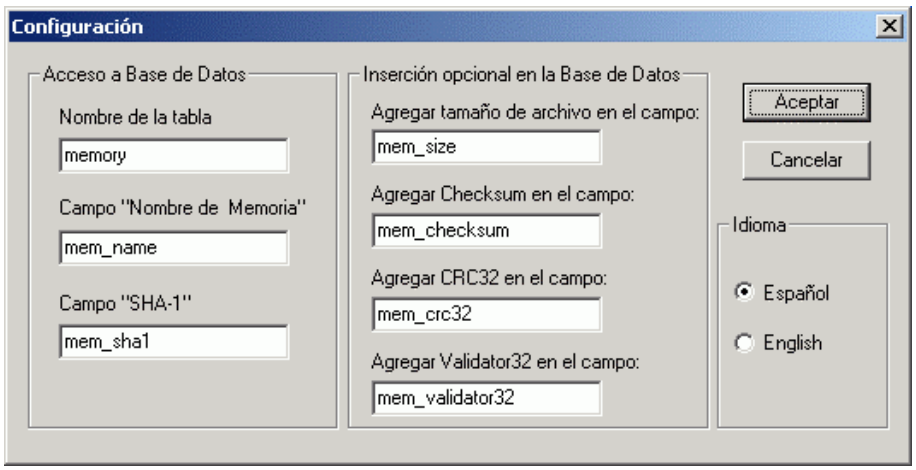


Figura 4. Ventana de configuración de la aplicación

Para realizar el cambio de idioma basta seleccionar entre las dos opciones presentes: Español o Inglés.

Una vez realizada la selección del idioma, el botón “Aceptar” revisará si ha habido un cambio en la selección. De ser el caso se anunciará la necesidad de reiniciar la aplicación para que los cambios tengan efecto.

3 Configuración para el acceso a bases de datos

DSC v3.1 ha sido diseñado para integrar el cálculo de huellas digitales de archivos con el acceso a bases de datos de huellas digitales que permitan precisar si la huella digital recién calculada se encuentra registrada en la base, lo que facilita determinar si el archivo analizado puede ser identificado como conocido (es decir, válido) o desconocido, lo cual puede ser interpretado como un archivo que no está registrado en la base de datos o un archivo registrado pero que al fallar en identificar la huella digital, se sospecha acerca de su autenticidad.

Para utilizar las características de acceso a bases de datos es necesario contar con los permisos necesarios de lectura (para comparación de huellas digitales) y escritura (para añadir registros a la base de datos). Esos permisos (usuarios y claves de acceso) deben ser entregados por el administrador de la base de datos.

La configuración del programa para realizar consultas e inserción de campos en la base de datos requiere información acerca del diseño de la misma. Consulte al administrador de su base de datos acerca de los nombres de tablas, campos y del diseño en general antes de intentar utilizar estas características de DSC.

3.1 Configurando la base de datos como Fuente de Datos ODBC

Debido a que existen muchos tipos de bases de datos (Desde una simple tabla separada por comas hasta sistemas completos de base de datos como Oracle, MS Access o MySQL) y a que el diseño de una base de datos puede ser tan variado como diseñadores de bases de datos existen, se han realizado las siguientes simplificaciones:

- El acceso a las bases de datos se realiza utilizando el servicio de Orígenes de Bases de Datos (ODBC) de los sistemas operativos Windows. A través de este mecanismo se puede conectar con cualquier motor de bases de datos que cuente con un software de conexión o "*driver ODBC*", con el cual se pueden formular

consultas al motor de la base de datos utilizando sentencias SQL estándar.

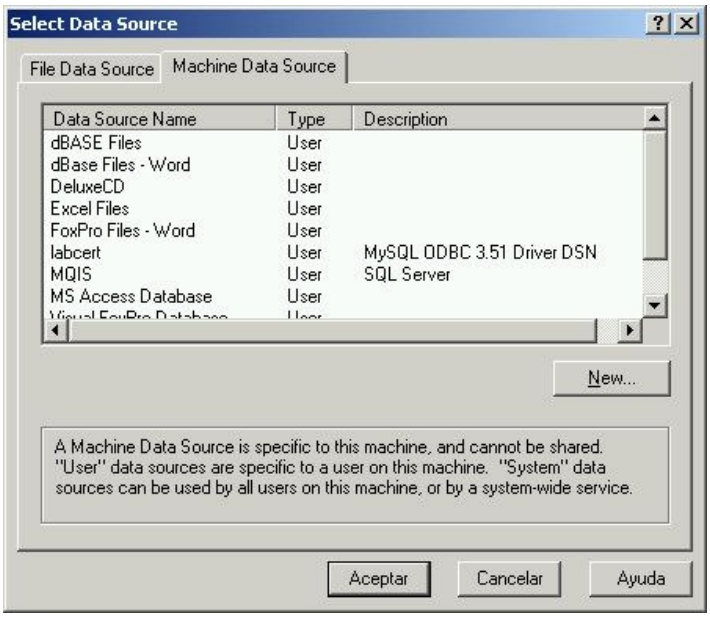


Figura 5. Selección de fuente de datos ODBC

- La administración de los permisos de acceso de lectura o lectura-escritura se realizan utilizando la interfase provista por cada software de conexión al motor de la base de datos.

DSC no interviene en la validación de los permisos de acceso a las bases de datos ni almacena dicha información en ninguna manera.

Consulte la información del software de conexión ODBC de su base de datos acerca de los parámetros que podrían ser requeridos para establecer su base de datos como una Fuente de Datos ODBC.

3.2 Configuración para ejecutar consultas (con permisos de lectura)

Una vez que se configura una base de datos como Fuente de Datos ODBC, es necesario contar con la siguiente información acerca del diseño de la base de datos:

- El nombre de la tabla dentro de la base de datos en donde se almacena la información sobre las huellas digitales.
- El nombre de los campos dentro de la tabla en donde se almacena la siguiente información:
  - El nombre registrado del archivo en la base de datos. Por defecto, este nombre es “mem_name”.
  - La huella digital SHA-1. Por defecto, este nombre es “mem_sha1”.

DSC realiza la consulta a la base de datos buscando registros que contengan el campo en donde se almacena la huella SHA-1 similar a la obtenida en la verificación de los archivos.

Con esta información es posible realizar la operación de “Compara con una base de datos”, luego de realizar la verificación de los archivos.

3.3 Configuración para la actualización de bases de datos (permisos de lectura y escritura)

Al insertar nuevos registros en la base de datos mediante la opción “Actualizar Base de Datos” es posible agregar los siguientes campos:

- El tamaño del archivo.
- El Checksum del archivo.
- El CRC-32 del archivo.
- El Validator-32 del archivo.

Estos campos serán agregados siempre y cuando se llene la casilla correspondiente con el nombre del campo que debe contener esta información. En la ventana de Configuración se muestran estos campos bajo el rubro “Inserción opcional en la Base de Datos”.

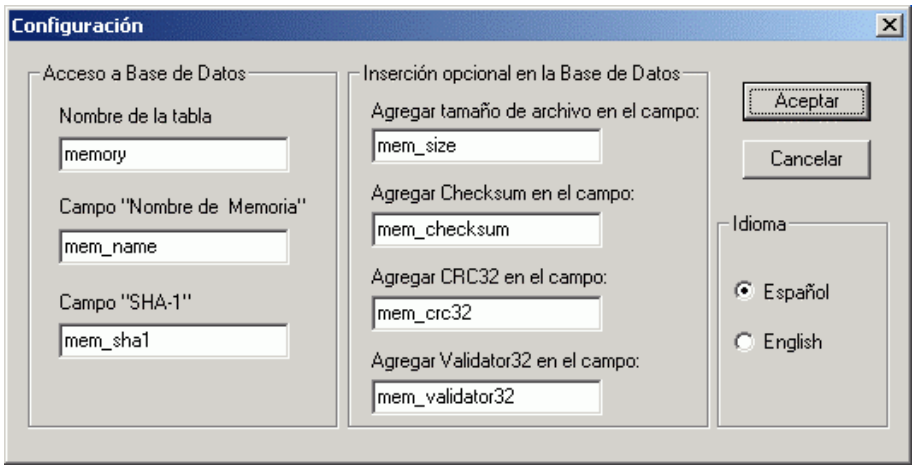


Figura 6. Campos opcionales para la inserción en la Base de Datos

Los nombres de la tabla y de los campos se almacenan en la memoria de la computadora, de manera que cada vez que se utiliza la aplicación se puede volver a interactuar con la base de datos sin necesidad de describir esta información nuevamente.

4 Comparando un resultado con una base de datos

Una vez que se ha realizado el cálculo de las huellas digitales de uno o varios archivos puede ser útil comparar la información obtenida con otra almacenada en una base de datos. Con esto, podremos determinar si los archivos que han sido almacenados están registrados en la base de datos o si la huella digital obtenida no es la esperada, con lo que se puede empezar a dudar de la validez del archivo analizado.

Para aprovechar esta característica incorporada en DSC es necesario seguir las instrucciones establecidas en el Capítulo 3 “Configuración para el acceso a Bases de Datos”.

Luego de calcular las huellas digitales de los archivos seleccionados, cada resultado es mostrado precedido de un símbolo de interrogación en un recuadro verde. Esto indica que las huellas digitales calculadas aun no han sido comparadas con ninguna fuente de datos.

Para proceder a la comparación, se deben realizar los siguientes pasos:

- Utilizar el botón “Comparar con Base de Datos”. Esta operación nos lleva al cuadro de diálogo del sistema operativo para la selección de una Fuente de Datos ODBC:

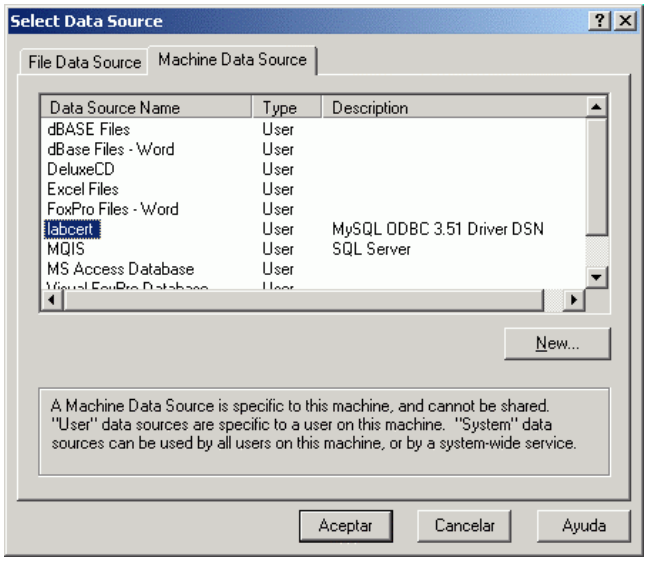


Figura 7. Selección de una Fuente de Datos ODBC

- Seleccione la Fuente de Datos ODBC que simboliza la base de datos a consultar. Dependiendo del motor de base de datos seleccionado y de su software de interfase, podría aparecer alguna ventana en donde se pueden configurar parámetros de acceso, como por ejemplo, nombres de usuario y claves de acceso a la base de datos. En la siguiente imagen se muestra la ventana de configuración del software de interfase del motor de base de datos MySQL:

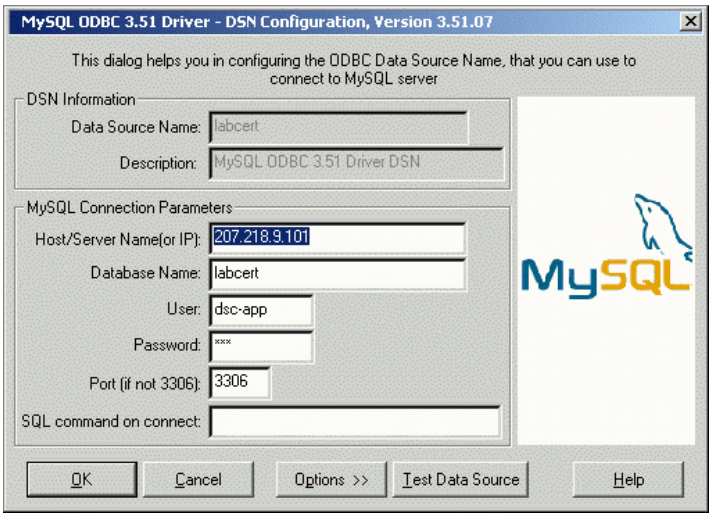


Figura 8. Ventana de acceso del controlador MySQL

Para utilizar esta característica es necesario que cuente al menos con una cuenta de usuario y clave de acceso con permisos de lectura de la tabla dentro de la base de datos.

- De contar con los permisos necesarios, DSC procede a emitir la consulta utilizando sentencias SQL estándar.

DSC realiza la consulta a la base de datos buscando registros que contengan el campo en donde se almacena la huella SHA-1, similar a la obtenida en la verificación de los archivos.

- Los resultados de la búsqueda en la base de datos se presentan de la siguiente manera:
  - Los archivos cuya huella digital SHA-1 coincide con el campo correspondiente de algún registro en la base de datos se presentan con un signo de confirmación (✓), en recuadro azul y se agrega la información encontrada en la base de datos para ese registro del nombre del archivo.

- Los archivos cuya huella digital SHA-1 no coincide con el campo correspondiente de ningún registro en la base de datos se presentan con un signo de negación (X), en recuadro rojo y la frase “NO EXISTE”.

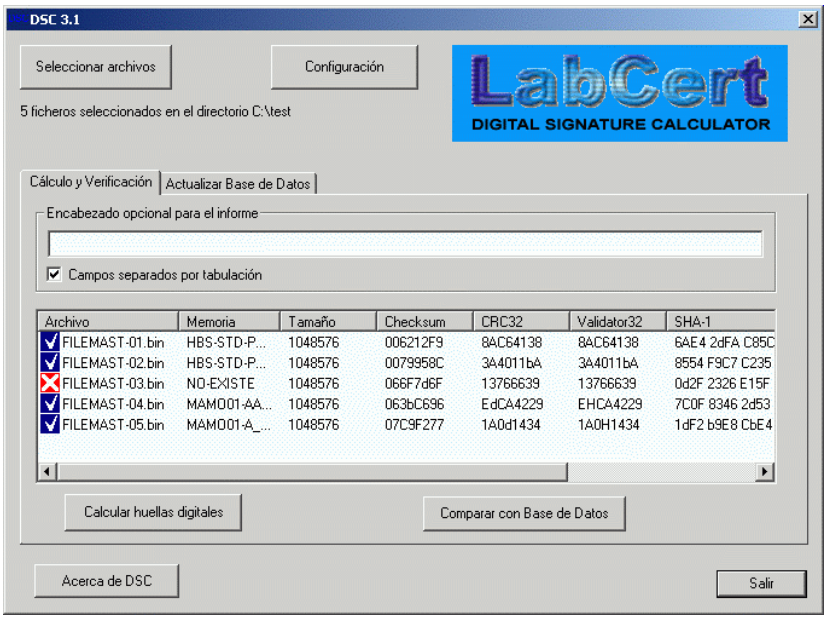


Figura 9. Comparación de resultados con una base de datos

DSC no adiciona o borra ningún registro ni altera ningún campo de la base de datos cuando se utiliza únicamente esta opción.

5 Agregando resultados a una base de datos

El proceso de alimentar una base de datos para después realizar consultas sobre los campos se facilita si se utiliza una única herramienta para ambas tareas.

DSC v3.1 ha sido diseñado para realizar ambas funciones. A través de una simple consulta SQL de inserción, se puede agregar nuevos registros a una base de datos que aloja las huellas digitales.

El acceso a bases de datos siempre involucra riesgos en la seguridad de una organización. DSC ofrece la funcionalidad de consulta a bases de datos utilizando el mecanismo ODBC de sistemas operativos Windows. Los controladores ODBC para diferentes motores de bases de datos proveen mecanismos de seguridad como nombres de usuario, claves de acceso y conexiones cifradas. Consulte al administrador de su base de datos acerca de los posibles riegos y parámetros de configuración necesarios. Consulte también el Capítulo 3 “Configuración para el acceso a Bases de Datos”.

5.1 Campos básicos que se agregan por registro

Cada archivo analizado genera la inserción de un registro en la tabla de la base de datos configurada. Los campos básicos que se agregan por registro son:

- Nombre del archivo.
- Huella digital SHA-1.

En diseños de bases de datos con campos automáticos se observará más información.

Si el diseño de la base de datos no permite registros vacíos que no se pueden llenar ninguna de estos datos, consulte en los temas “Campos opcionales” y “Campos adicionales”.

5.2 Campos opcionales

En cada registro añadido, se puede llenar campos con la siguiente información:

- Tamaño del archivo.
- Checksum del archivo.
- CRC-32 del archivo.
- Validator32 del archivo.

5.3 Campos adicionales

En cada registro añadido se puede llenar hasta dos campos adicionales con información variada, que puede provenir de las siguientes fuentes:

- Texto estático. Todos los registros añadidos tendrán en el nombre de campo, el texto estático especificado.
- Valor del registro de otra tabla de datos. Involucra el acceso a una tercera fuente de información (otra base de datos). Para este acceso hay que repetir todo el mecanismo de validación y acceso a una base de datos reconocida como Fuente de Datos ODBC.

En la opción “Actualizar Base de Datos” de DSC se muestra el área de configuración para obtener la información para los dos campos adicionales que se agregarán en cada registro.

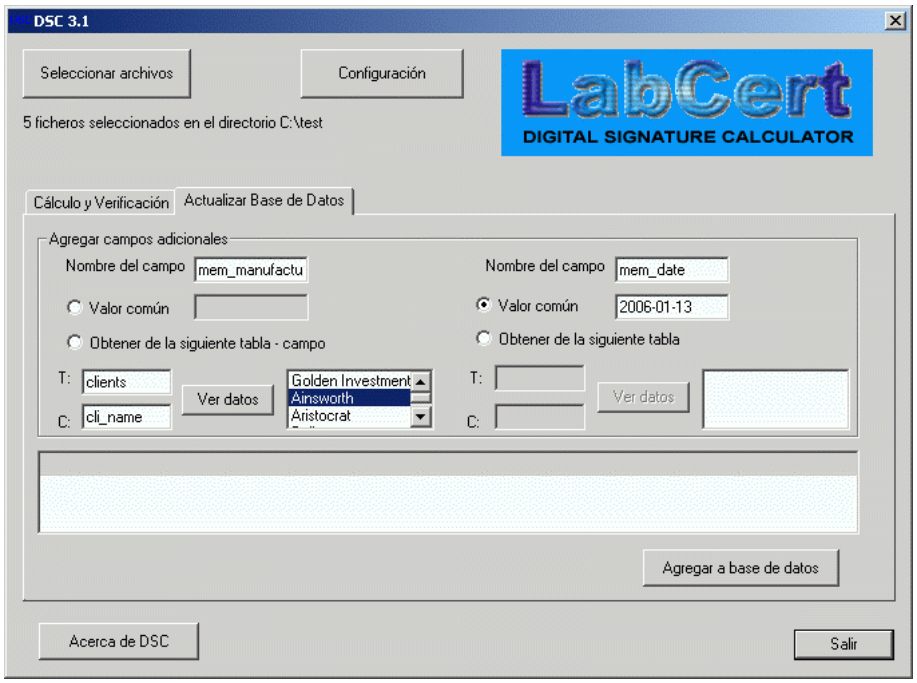


Figura 10. Actualización de base de datos

Para cada uno de los dos campos adicionales se configura lo siguiente:

- Nombre del campo: Es el nombre del campo dentro de la tabla a actualizar.
- Opciones para el llenado del campo:
  - Valor común: involucra el uso de un texto estático, que se agregará a cada registro.
  - Obtener de la siguiente tabla-campo: requiere agregar el nombre de una tabla (T) y de un campo (C). Al proveer esta información y presionar el botón “Ver datos” se abrirá la ventana de diálogo de Fuentes de Datos ODBC para seleccionar la fuente de datos apropiada (involucra también poseer la información necesaria de parámetros y configuraciones, como nombres de usuarios y contraseñas). En la lista contigua aparecerán los datos extraídos de esta fuente de datos, para seleccionar aquel que será utilizado en la operación de inserción de registros.