Consejería de Hacienda y Administración Pública

Módulo de firmas PDF

Sevilla, noviembre de 2010

Hoja de control		
Fecha	**Autor**	**Descripción**
03/11/2010	RPV	Adaptación módulo de firmas PDF MPR

1 Introducción

El Cliente de Firma es una herramienta de Firma Electrónica que funciona en forma de Applet de Java integrado en una página Web mediante JavaScript.

El Cliente hace uso de los certificados digitales X.509 y de las claves privadas asociadas a los mismos que estén instalados en el repositorio o almacén de claves y certificados (*keystore*) del navegador web (*Internet Explorer, Mozilla, Firefox*) o el sistema operativo así como de los que estén en dispositivos (tarjetas inteligentes, dispositivos *USB*) configurados en el mismo (el caso de los DNI-e).

El Cliente de Firma, como su nombre indica, es una aplicación que se ejecuta en cliente (en el ordenador del usuario, no en el servidor Web). Esto es así para evitar que la clave privada asociada a un certificado tenga que "salir" del contenedor del usuario (tarjeta, dispositivo USB o navegador) ubicado en su PC. De hecho, nunca llega a salir del navegador, el Cliente le envía los datos a firmar y éste los devuelve firmados.

El Cliente de Firma contiene las interfaces y componentes web necesarios para la realización de los siguientes procesos (además de otros auxiliares como cálculos de hash, lectura de ficheros, etc…):

- Firma de formularios Web.
- Firma de datos y ficheros.
- Multifirma masiva de datos y ficheros.
- Cofirma (CoSignature) → Multifirma al mismo nivel.
- Contrafirma (CounterSignature) → Multifirma en cascada.

Como complemento al cliente de firma, se encuentra un cliente de cifrado que nos permite realizar las funciones de encriptación y desencriptación de datos atendiendo a diferentes algoritmos y configuraciones. Además permite la generación de sobres digitales.

2 Objetivos

El objetivo del presente documento es reflejar las capacidades del Cliente @firma en cuanto a firmas electrónicas de documentos PDF (*Portable Document Format*), las peculiaridades de este relativas a este formato, sus opciones de configuración y su compatibilidad.

3 Formato de firma PDF (Portable Document Format)

El cliente soporta firmas en formato PDF / ISO 32000-1 mediante algoritmo RSA y huellas digitales según algoritmos SHA-1 y SHA-2, de forma acorde a las especificaciones de Adobe. Este tipo de firma consiste, básicamente, en la inclusión de una estructura de firma CAdES en un espacio reservado dentro del documento PDF que se firma.

Las firmas de los documentos PDF realizadas por el cliente **no son visibles directamente**, es decir, no tienen representación gráfica dentro del documento.

Soporte de firmas PDF Avanzadas (PAdES)

La versión actual del Cliente @firma genera firmas electrónicas compatibles con la especificación PAdES (http://pda.etsi.org/pda/) de un modo completamente compatible con la versión 9.3.2 de Adobe Reader y con la normativa ISO 32000-1. Todas las firmas PDF realizadas por el Cliente son simultáneamente compatibles con la normativa ISO 32000-1 y el estándar PAdES.

Las firmas electrónicas introducidas pueden generarse como PAdES-BES (firmas avanzadas básicas) o PAdES-EPES (firmas avanzadas acordes a política), siendo necesario en este último caso introducir los datos de la política de firmas tal y como se describe en la documentación general del Cliente @firma (JavaDoc del Applet y Manual del Integrador):

```
setPolicy(String identifier, String description, String qualifier)
```

A nivel técnico, los datos de firma electrónica empotrados dentro de la estructura PDF equivalen a una firma **CAdES** con **Signing Certificate V2**.

El Cliente no soporta la firma de adjuntos a los documentos PDF en ninguna variante ni formato.

3.1 Uso de las firmas PDF desde el cliente @firma

El formato de firma PDF está únicamente disponible en la construcción Completa del cliente de firma.

Para realizar una firma en este formato sólo es necesario configurarlo como formato de firma mediante el método del Applet cliente:

```
setSignatureFormat(String format)
```

Puede llamarse a esta función desde las páginas Web que integren el cliente por medio de la sentencia JavaScript:

```
clienteFirma.setSignatureFormat(String format)
```

Esto configurará el cliente de firma para realizar firmas PDF, siendo necesario también completar la configuración del cliente que se considere necesaria (algoritmo de firma, tratamiento de errores,…), establecer los datos que se desean firmar (no es posible realizar firmas PDF a partir del hash de un documento) y ordenar el proceso de firma.

3.2 Parámetros de funcionamiento

- Cadena de identificación de formato (dos variantes aceptadas, se ignoran mayúsculas y minúsculas):
  - "PDF" / "Adobe PDF"
- Se ignorará cualquier indicación del modo de firma (implícito o explícito).
- El formato PDF no es compatible con el algoritmo de firma MD2withRSA.
- Ficheros de entrada
  - Adobe PDF / ISO 32000 (*.pdf)
- Ficheros de salida
  - Adobe PDF / ISO 32000 (*.pdf)

A modo de ejemplo, la llamada para establecer el modo de funcionamiento de firmas PDF se usaría la siguiente llamada:

```
clienteFirma.setSignatureFormat("PDF");
```

3.2.1 Co-Firmas en PDF (Firmas en serie)

Los documentos PDF pueden ser multi-firmados mediante el módulo, pero debe observarse la siguientes peculiaridades:

- Las co-firmas PDF son en realidad múltiples firmas aplicadas sobre un mismo documento sobre el que se crean revisiones, con el siguiente funcionamiento:
- La primera firma aplicada aplica al documento original.
- La segunda firma crea una nueva revisión del documento, que contiene el documento original **más la primera firma**, y esta la revisión firmada.

- Cada firma adicional crea una nueva revisión que contiene tanto el documento original como las firmas anteriores.
- Una co-firma PDF, responde más al concepto de documento "multifirmado" que al concepto de co-firma aplicado a firmas CMS o XMLdSig, ya que internamente el PDF no contiene una única estructura PKCS#7 con las firmas / co-firmas, sino varias estructuras PKCS#7.
- El formato PDF permite, tras una firma digital, sellar el documento para evitar que se añada contenido adicional tras la firma. Dado que esto evitaría la posibilidad de añadir nuevas firmas, **el módulo no sella los documentos, posibilitando la anexión de nuevo contenido al documento.**
- No obstante, la primera firma aplicará únicamente al contenido original, aspecto que se indicará (por el programa Adobe Reader) como advertencia. Este aspecto en absoluto invalida la firma.
- El formato PDF permite reservar espacio para firmas de modo que no sea necesario anexar contenido al añadir firmas digitales adicionales (tras una primera firma), ya que estas se alojan en estos espacios reservados. Pero dado que es necesario determinar a priori el número de espacios a reservar (uno por cada firma digital adicional a la primera) y que el cliente carece de esta información, no se reserva ningún espacio. Cada firma adicional a la primera supone un contenido anexado al documento.
- Al ser las firmas adicionales a la primera contenido anexo, no existente en el momento de realizar la primera operación de firma, esta no las reflejará, **indicando mediante una advertencia que se ha añadido contenido adicional**, pero igualmente informando que el contenido firmado no ha sufrido alteraciones.
- El cliente no puede co-firmar documentos PDF sellados.
- No es posible realizar co-firmas a partir de un hash.

3.2.2 Contrafirmas en PDF

El módulo de firmas PDF no soporta contrafirmas.

3.3 Información de utilidad

- Información sobre el formato PDF
  - http://www.adobe.com
- Normativa PKCS#7
  - http://www.rsa.com/rsalabs/node.asp?id=2129

4 Glosario de términos

Firma electrónica

Es el conjunto de datos, en forma electrónica, anejos a otros datos electrónicos o asociados funcionalmente con ellos, utilizados como medio para identificar formalmente al autor o a los autores del documento que la recoge.

XML Digital Signature (XMLDSig)

Es una recomendación del W3C que define una sintaxis XML para la firma digital

XML Advanced Signature (XAdES)

Es un conjunto de extensiones a las recomendaciones XML-DSig haciéndolas adecuadas para la firma electrónica avanzada.

RSA

Es un sistema criptográfico de clave pública desarrollado en 1977. En la actualidad, RSA es el primer y más utilizado algoritmo de este tipo y es válido tanto para cifrar como para firmar digitalmente.

XML

Es un metalenguaje extensible de etiquetas desarrollado por el World Wide Web Consortium (W3C). Es una simplificación y adaptación del SGML y permite definir la gramática de lenguajes específicos (de la misma manera que HTML es a su vez un lenguaje definido por SGML). Por lo tanto XML no es realmente un lenguaje en particular, sino una manera de definir lenguajes para diferentes necesidades. Algunos de estos lenguajes que usan XML para su definición son XHTML, SVG, MathML.

Office Open XML (OOXML)

Es un formato de archivo abierto y estándar cuyas extensiones más comunes son .docx, .xlsx y .pptx. Se le utiliza para representar y almacenar hojas de cálculo, diagramas, presentaciones y documentos de texto. Un archivo Office Open XML contiene principalmente datos basados en el lenguaje de marcado XML, comprimidos en un contenedor .zip específico.

Open Document Format (ODF)

Es un formato de fichero estándar para el almacenamiento de documentos ofimáticos tales como hojas de cálculo, memorandos, gráficas y presentaciones. Aunque las especificaciones fueron inicialmente elaboradas por Sun, el estándar fue desarrollado por el comité técnico para Open Office XML de la organización OASIS y está basado en un esquema XML inicialmente creado e implementado por la suite ofimática OpenOffice.org (ver OpenOffice.org XML).

ZIP

Es un formato de almacenamiento sin pérdida, muy utilizado para la compresión de datos como imágenes, programas o documentos.

PDF

Es un formato de almacenamiento de documentos, desarrollado por la empresa Adobe Systems. Este formato es de tipo compuesto (imagen vectorial, mapa de bits y texto).

SHA

Es un sistema de funciones hash criptográficas relacionadas de la Agencia de Seguridad Nacional de los Estados Unidos y publicadas por el National Institute of Standards and Technology (NIST). El primer miembro de la familia fue publicado en 1993 es oficialmente llamado SHA. Sin embargo, hoy día, no oficialmente se le llama SHA-0 para evitar confusiones con sus sucesores. Dos años más tarde el primer sucesor de SHA fue publicado con el nombre de SHA-1. Existen cuatro variantes más que se han publicado desde entonces cuyas diferencias se basan en un diseño algo modificado y rangos de salida incrementados: SHA-224, SHA-256, SHA-384, y SHA-512 (llamándose SHA-2 a todos ellos).

PKCS

Se refiere a un grupo de estándares de criptografía de clave pública concebidos y publicados por los laboratorios de RSA en California. A RSA Security se le asignaron los derechos de licenciamiento para la patente de algoritmo de clave asimétrica RSA y adquirió los derechos de licenciamiento para muchas otras patentes de claves.

W3C

Es un consorcio internacional que produce recomendaciones para la World Wide Web. Está dirigida por Tim Berners-Lee, el creador original de URL (Uniform Resource Locator, Localizador Uniforme de Recursos), HTTP (HyperText Transfer Protocol, Protocolo de Transferencia de HiperTexto) y HTML (Lenguaje de Marcado de HiperTexto) que son las principales tecnologías sobre las que se basa la Web.

OpenOffice.org

es una suite ofimática libre (código abierto y distribución gratuita) que incluye herramientas como procesador de textos, hoja de cálculo, presentaciones, herramientas para el dibujo vectorial y base de datos. Está disponible para varias plataformas, tales como Microsoft Windows, GNU/Linux, BSD, Solaris y Mac OS X. Soporta numerosos formatos de archivo, incluyendo como predeterminado el formato estándar ISO/IEC OpenDocument (ODF), entre otros formatos comunes. A febrero de 2010, OpenOffice soporta más de 110 idiomas.

Base64

Es un sistema de numeración posicional que usa 64 como base. Es la mayor potencia de dos que puede ser representada usando únicamente los caracteres imprimibles de ASCII. Esto ha propiciado su uso para codificación de correos electrónicos, PGP y otras aplicaciones. Todas las variantes famosas que se conocen con el nombre de Base64 usan el rango de caracteres A-Z, a-z y 0-9 en este orden para los primeros 62 dígitos, pero los símbolos escogidos para los últimos dos dígitos varían considerablemente de unas a otras. Otros métodos de codificación como UUEncode y las últimas versiones de binhex usan un conjunto diferente de 64 caracteres para representar 6 dígitos binarios, pero éstos nunca son llamados Base64.

ASN.1

Es una norma para representar datos independientemente de la máquina que se esté usando y sus formas de representación internas. Es un protocolo de nivel de presentación en el modelo OSI.

Autoridad de Certificación (CA)

Es una entidad de confianza, responsable de emitir y revocar los certificados digitales o certificados, utilizados en la firma electrónica, para lo cual se emplea la criptografía de clave pública. Jurídicamente es un caso particular de Prestador de Servicios de Certificación.

Certificado Digital

Es un documento digital mediante el cual un tercero confiable (una autoridad de certificación) garantiza la vinculación entre la identidad de un sujeto o entidad y su clave pública.

Infraestructura de Clave Pública (PKI)

Es una combinación de hardware y software, políticas y procedimientos de seguridad que permiten la ejecución con garantías de operaciones criptográficas como el cifrado, la firma digital o el no repudio de transacciones electrónicas.

5 Información de contacto

Soporte a la Administración Electrónica
Consejería de Hacienda y Administración Pública
Dirección General de Tecnologías para Hacienda y la Administración Electrónica
soporte.admonelectronica@juntadeandalucia.es