Cálculo de la huella digital de documentos electrónicos

Criterios para el cálculo del hash

Documento de Integración
Sistemas Desarrollo
Versión
001
Fecha de revisión
27/09/2022
Realizado por
Servicio de Firma y Gestión Documental

Contenido

1. Control de modificaciones

1.1 Revisión actual

Revisión: 001

Fecha: 27/09/2022

Autor: Servicio de Firma y Gestión Documental

- Versión inicial del documento

2. Introducción

La agrupación de un conjunto de documentos electrónicos se realiza en el expediente electrónico. Este expediente electrónico consiste en **un índice que lista todos los documentos** que forman parte de él. Tanto los documentos como los expedientes electrónicos siguen el formato ENI.

Como medida de control, **para asegurar que el documento referenciado en el índice del expediente y el documento enviado junto a él es el mismo, se realiza el cálculo de un hash** (una secuencia de caracteres) de la firma/contenido del documento electrónico y **este hash se añade al índice del expediente** junto a la referencia del documento como su huella digital.

El cálculo de un hash se puede realizar mediante diversos algoritmos, de manera que al aplicar el mismo algoritmo a un contenido siempre se obtiene el mismo hash (la misma huella digital).

Por tanto, también **se añade el algoritmo con el que se ha calculado el hash al índice del expediente** junto a la referencia del documento.

Esta es la manera de asegurar la integridad de todos los documentos enviados y confiar en que ninguno de ellos ha sido manipulado.

Los algoritmos de cálculo de hash admitidos son los siguientes:

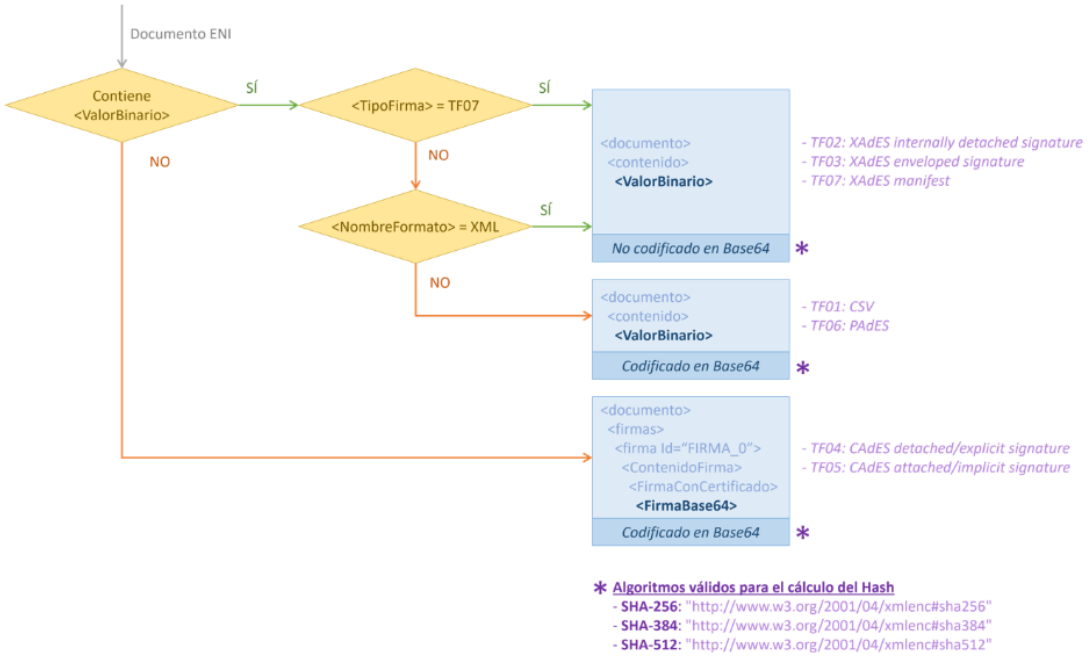
- SHA-256
- SHA-384
- SHA-512 (recomendamos utilizar este).

3. Obtención del contenido con el que calcular la huella

Dependiendo de la firma del documento ENI, el contenido a utilizar para calcular el hash variará, así como si este contenido debe estar codificado en Base64 o no.

Nunca debe calcularse el hash del documento ENI completo.

El siguiente gráfico representa el flujo a seguir para conocer el contenido con el que calcular las huellas de los documentos e incluirlas en el índice del expediente ENI.



Se distinguen 4 casos según el tipo de firma o formato del fichero del documento ENI:

- **Caso A:** El Documento ENI contiene la etiqueta <ValorBinario> y se encuentra una etiqueta <TipoFirma> con el valor "TF07".

  El elemento <ValorBinario> se encuentra dentro de la etiqueta <contenido> del Documento ENI, como se puede ver en el ejemplo:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<insidews:documento xmlns:ns2="http://administracionelectronica.gob.es/ENI,
    <ns5:documento Id="ES_E04975701_2022_8951">
        <contenido Id="CONTENIDO_DOCUMENTO">
            <ValorBinario>JVBERi0xLjcKJeLjz9MKMSAwIG9iago8PC9UeXBlL1NpZy9G
            <NombreFormato>PDF</NombreFormato>
        </contenido>
```

  El contenido a utilizar para calcular la huella es **el valor de la etiqueta <ValorBinario> decodificado en Base64**.

- **Caso B:** El Documento ENI contiene la etiqueta <ValorBinario>, no existe ninguna etiqueta <TipoFirma> con el valor "TF07" y la etiqueta <NombreFormato> tiene el valor "XML".

  La etiqueta <NombreFormato> se encuentra dentro de la etiqueta <contenido>, junto a la etiqueta <ValorBinario>.

  El contenido a utilizar para calcular la huella es **el valor de la etiqueta <ValorBinario> decodificado en Base64**.

- **Caso C:** El Documento ENI contiene la etiqueta <ValorBinario>, no existe ninguna etiqueta <TipoFirma> con el valor "TF07" y la etiqueta <NombreFormato> NO tiene un valor diferente a "XML".

  La etiqueta <NombreFormato> se encuentra dentro de la etiqueta <contenido>, junto a la etiqueta <ValorBinario>.

  El contenido a utilizar para calcular la huella es **el valor de la etiqueta <ValorBinario>** (que ya viene codificado en Base64).

- **Caso D:** El Documento ENI NO contiene el elemento <ValorBinario>.

  En su lugar debe venir la etiqueta <referenciaFichero>, tal como se muestra en este ejemplo:

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<insidews:documento xmlns:ns2="http://administracionelectronica.gob
    <ns5:documento Id="ES_E04975701_2022_8950">
        <contenido Id="CONTENIDO_DOCUMENTO">
            <referenciaFichero>#FIRMA_0</referenciaFichero>
            <NombreFormato>DOCX</NombreFormato>
        </contenido>
```

  El valor de la etiqueta <referenciaFichero> es el identificador de la firma con el contenido del fichero.

  Dentro de la etiqueta <firmas> se debe buscar una etiqueta <firma> cuyo "Id" sea el valor de <referenciaFichero> (quitando el carácter de la almohadilla '#'), como se muestra en este ejemplo:

```
        <ns3:firmas>
            <ns3:firma Id="FIRMA_0">
                <ns3:TipoFirma>TF05</ns3:TipoFirma>
                <ns3:ContenidoFirma>
                    <ns3:FirmaConCertificado>
                        <ns3:FirmaBase64>MIJHPAYJKoZIhvcNAQ
                    </ns3:FirmaConCertificado>
                </ns3:ContenidoFirma>
            </ns3:firma>
        </ns3:firmas>
```

  El contenido a utilizar para calcular la huella es **el valor de la etiqueta <FirmaBase64>** (que ya viene codificado en Base64).

4. Traslado de la huella y el algoritmo al expediente

Una vez obtenido el contenido sobre el que obtener el hash se le aplicará el algoritmo elegido.

Los algoritmos de cálculo de Hash admitidos son los siguientes:

- SHA-256
- SHA-384
- SHA-512 (recomendamos utilizar este).

Tras aplicar el algoritmo al contenido se obtendrá finalmente el hash (huella digital).

En el índice del expediente ENI, junto al identificador ENI de cada documento se debe informar del algoritmo elegido para obtener el hash (etiqueta <FuncionResumen>), así como del propio hash obtenido (etiqueta <ValorHuella>), tal como se muestra en el ejemplo:

```
<ns2:indice Id="EXP_INDICE_ES_E04975701_2022_EXP_6078">
    <ns2:IndiceContenido Id="EXP_INDICE_CONTENIDOES_E04975701_2022_EXP_6078">
        <FechaIndiceElectronico>2022-09-12T10:22:47.000+02:00</FechaIndiceElectronico>
        <DocumentoIndizado>
            <IdentificadorDocumento>ES_E04975701_2022_8879</IdentificadorDocumento>
            <ValorHuella>c27b1571d7ef025254d2764bf1390fd3d3ffb21f15b64bd507c698f045b5fe4df399561acc01d51
            <FuncionResumen>http://www.w3.org/2001/04/xmlenc#sha512</FuncionResumen>
            <FechaIncorporacionExpediente>2022-09-12T00:00:00.000+02:00</FechaIncorporacionExpediente>
            <OrdenDocumentoExpediente>00001</OrdenDocumentoExpediente>
        </DocumentoIndizado>
    </ns2:IndiceContenido>
```

El valor de la etiqueta <FuncionResumen> según el algoritmo aplicado será:

- **Para SHA-256**: http://www.w3.org/2001/04/xmlenc#sha256
- **Para SHA-384**: http://www.w3.org/2001/04/xmlenc#sha384
- **Para SHA-512**: http://www.w3.org/2001/04/xmlenc#sha512