(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 802 448**

(51) Int. CI.:

***H04N 1/32*** (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **30.03.2009 E 09156764 (4)**

(97) Fecha y número de publicación de la concesión europea: **13.05.2020 EP 2237546**

(54) Título: **Dispositivo y proceso para proteger un documento digital, y proceso correspondiente para verificar la autenticidad de una copia impresa**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
**19.01.2021**

(73) Titular/es:

**CIPAL SCHAUBROECK NV (100.0%)**
**Cipalstraat 3**
**2440 Geel, BE**

(72) Inventor/es:

**BALS, KLAAS;**
**DEHOND, GUY y**
**HOFSTEDE, NICK**

(74) Agente/Representante:

**SÁEZ MAESO, Ana**

## DESCRIPCIÓN

Dispositivo y proceso para proteger un documento digital, y proceso correspondiente para verificar la autenticidad de una copia impresa

Campo de la invención

La presente invención generalmente se refiere a proteger la autenticidad de copias digitales y copias impresas (copias analógicas o versiones impresas en papel) de un documento. Con un estimado de 250 000 casos de fraudes de documentos oficiales/año, se estima que un tercio de todos los supuestos documentos oficiales del gobierno son falsos, y un costo comunitario estimado de 40 000 euros por documento falsificado, la necesidad de una solución de gran eficiencia que garantice la autenticidad de los documentos que existen tanto en forma digital como analógica se han vuelto enormes. Ejemplos de documentos que están sujetos a fraude y donde existe la necesidad de verificar la autenticidad de las copias digitales y analógicas son los extractos bancarios, recibos de pago, documentos gubernamentales, contratos temporales tales como contratos provisionales, certificados de origen que acompañan a envíos internacionales de mercancías, facturas, certificados utilizados en entornos altamente seguros tales como defensa, nuclear, aeroespacial, comercio de diamantes, etc. El objetivo general de la presente invención es proporcionar una solución para proteger de manera eficiente la autenticidad de las copias digitales y analógicas de documentos, para así estimular el uso y la difusión de documentos digitales certificados al tiempo que permite a las personas imprimir sus documentos digitales mediante impresoras comunes en papel común con tinta común con garantía del valor legal y autenticidad de los documentos impresos.

Antecedentes de la Invención

Aunque los documentos digitales son rentables en generación, multiplicación y entrega, los documentos en papel tienen una larga historia cultural y siguen siendo el medio preferido para el mantenimiento de registros. Sin embargo, producir y distribuir copias en papel es más costoso. Los documentos en papel se pueden replicar mediante fotocopias y se pueden distribuir en forma original o fotocopiada, por ejemplo, por fax o correo postal.

Mientras que los documentos digitales pueden tener firmas digitales que prueban la autenticidad, la firma digital se pierde cuando los documentos se imprimen como consecuencia de lo cual ya no se puede probar la autenticidad de las versiones en papel. Esto da como resultado un riesgo de falsificación o fraude de documentos.

La autenticidad es un tema importante para documentos oficiales tales como tarjetas de identificación o pasaportes. Otros documentos para los cuales se desea resistencia a la falsificación u otra manipulación son varios certificados, licencias de conducir, boletas de apuestas, premios de juegos, boletos, extractos bancarios, certificados de origen, recibos de sueldo, formularios gubernamentales, documentos oficiales, documentos aduaneros, pólizas de seguros u otros documentos que simplemente requieran validar firmas adheridas a los mismos, tales como contratos, etc.

En resumen, sería muy ventajoso que los documentos sensibles, tales como los instrumentos negociables, se puedan generar a pedido a partir de una copia digital certificada sin requerir suministros especiales de papel de antecedentes de seguridad preimpresos, al tiempo que se garantiza que el documento impreso no se falsifique o cambie.

Se han descrito varios métodos en la literatura que intentan detectar u obstaculizar la falsificación o alteración de documentos sensibles. Sin embargo, las soluciones existentes de la técnica anterior que intentan abordar el problema general anterior no funcionan correctamente o, en caso de que funcionen, carecen de eficiencia, es decir, requieren mucho tiempo y/o requieren la inclusión de códigos largos aplicados visualmente en los documentos impresos. Las soluciones más cercanas conocidas de la técnica anterior y sus respectivas deficiencias se analizan en los siguientes párrafos.

La solicitud de patente de Estados Unidos núm. 2006/0271787 A1 (DeYoung y otros) titulada "System and Method for Validating a Hard-Copy Document Against an Electronic Version" describe un método para verificar la autenticidad de una copia impresa mediante la colocación de un código único llamado "firma digital" en US 2006/0271787 A1 al documento. La "firma digital" puede ser, por ejemplo, un código de barras 2-D o Glossmarks™ como se indica en el documento US 2006/0271787 A1. En el código único, se codifica un resumen del mensaje del documento. Este resumen del mensaje se genera utilizando un algoritmo hash unidireccional en el contenido del documento como se especifica en el documento US 2006/0271787 A1. La verificación de la autenticidad de una copia impresa del documento requiere escanear la copia impresa y reproducir el resumen de mensaje o hash del documento escaneado. El hash reproducido se puede comparar con el resumen del mensaje o el hash que se obtiene decodificando el código de barras 2-D o la "firma digital" para detectar el fraude.

Un gran problema con el documento US 2006/0271787 A1 es que la copia impresa escaneada en general no permite la reproducción del hash. Como resultado de la resolución de escaneo limitada, el ruido aleatorio del escáner, las manchas o pequeñas manchas resultantes de la suciedad, etc., la copia impresa escaneada diferirá de la copia impresa y, en consecuencia, también del documento digital original. Incluso una pequeña diferencia en la copia impresa escaneada dará como resultado un hash significativamente diferente como resultado de lo cual la prueba de autenticación fallará, incluso si el documento no fue alterado.

Una solución equivalente de la técnica anterior en donde un hash (resumen del mensaje 24) de datos del documento (mensaje 20) se codifica de forma legible por máquina (firma digital 30) y se imprime en el documento para permitir la verificación de la autenticidad de la copia impresa se describe en la solicitud de patente de Estados Unidos núm. 2006/0265590 A1 titulada "Digital Signature/Certificate for Hard-Copy Documents" y mencionando al mismo co-inventor DeYoung. Como se ilustra en la Figura 1 y la Figura 3 del documento US 2006/0265590 A1, la verificación de la autenticidad requiere escanear la copia impresa y obtener el hash del documento escaneado para compararlo con el hash codificado en la firma digital. Sin embargo, el documento escaneado contendrá manchas de ruido y manchas como resultado de defectos de suciedad, escaneo e impresión, como resultado de lo cual el hash no se puede reproducir y ninguno o pocos documentos pasarán la prueba de autenticación con éxito. Además, al confiar en el texto extraído u otros datos en lugar de la imagen escaneada directamente, este método solo es factible cuando el documento se crea siguiendo reglas estrictas. Por ejemplo, el reconocimiento óptico de caracteres (OCR) con el nivel de precisión perfecto requerido solo es posible con fuentes lo suficientemente grandes y un tipo de letra adecuado. Otros métodos de extracción de datos conocidos en la técnica tienen inconvenientes similares.

También en el documento US 6,081,610 A *titulado "System and Method for Verifying Signatures on Documents",* solo se incluye un hash firmado en el código bidimensional (por ejemplo, código de barras) que se registra en el documento. Los diferentes equipos de escáner, la colocación ligeramente diferente del documento debajo del escáner, las arrugas en el documento, etc., pueden dar lugar a la reproducción de un hash completamente diferente del documento escaneado, de manera que la autenticidad del documento ya no se pueda verificar.

En otra solución de la técnica anterior, conocida de *Xerox Corporation's* de la patente de Estados Unidos 7,197,644 B2 *con el título "Systems and Methods for Providing Hardcopy Secure Documents and for Validation of such Documents" o su correspondiente solicitud de patente europea* EP 1 432 234 A1, una plantilla se utiliza para seleccionar un segmento de imagen, por ejemplo, en cheques o billetes de banco. El segmento de imagen y la plantilla están opcionalmente encriptados (Col. 3, líneas 30-35) y posteriormente codificados en, por ejemplo, un código holográfico, código de banda magnética, código de barras de alta densidad, código de microdot, código de glifo de datos, etc. (Col. 11, líneas 26-38). En el proceso de validación descrito en la columna 2, líneas 50-67 del documento US 7,197,644, se escanea el documento impreso que contiene la firma de imagen codificada y la plantilla de firma de imagen. La firma de imagen y la plantilla de firma de imagen se obtienen por decodificación (y descifrado opcional). El documento escaneado está sujeto a procesamiento adicional de acuerdo con la plantilla de firma de imagen para identificar en él el segmento de imagen para compararlo con el segmento de imagen en la firma de imagen decodificada.

Aunque el método del documento US 7,197,644 B2 no tiene que reproducir un hash de documentos escaneados como el documento US 2006/0265590 A1 y por lo tanto podría funcionar, solo opcionalmente cifra el segmento de imagen. El cifrado asimétrico del segmento de imagen, como se sugiere opcionalmente en el documento US 7,197,644 B2, tiene la desventaja de que esto consume mucho tiempo y es ineficiente.

Inconvenientes adicionales del documento US 7,197,644 B2 incluyen el hecho de que los datos de imagen, es decir, un subconjunto de la representación de mapa de bits del documento original, deben usarse para generar el código. Como resultado, el código generado será largo.

Otro inconveniente del documento US 7,197,644 B2 es que requiere que la plantilla de imagen se codifique en el código. La plantilla de imagen puede ser, por ejemplo, una franja rectilínea, una franja curva, un patrón de tablero de ajedrez y similares. El requisito de integrar la plantilla en el código aumentará aún más la longitud del código y contribuirá a la complejidad del sistema.

Todavía se conoce otra solución de la técnica anterior de IBM y se describe en la *solicitud de patente europea* EP 0 676 877 *titulada "Method and apparatus for authentication and verification of printed documents using digital signatures and authentication codes".* El método conocido de IBM adolece de los inconvenientes de la técnica anterior de DeYoung citada anteriormente (el escaneo introduce defectos como resultado de los cuales falla la verificación de la autenticidad) y el inconveniente adicional de la técnica anterior de Xerox citada anteriormente para segmentos figurativos (se generan códigos largos). Aunque IBM reconoce el problema de la resolución de escaneo limitada, el ruido del escáner, las manchas, etc. y se propone una solución para hacer frente a este problema, el método de IBM sigue siendo complejo y sensible a los errores, como se explicará en los siguientes párrafos.

El método de IBM para autenticar un documento comienza desde un documento en papel original que se escanea con un escáner convencional (etapa 1 en la página 4, líneas 7-8). En el método de IBM, el documento en papel escaneado se segmenta (etapa 2 en la página 4, líneas 9-11) de tal manera que diferentes segmentos contienen diferentes tipos de datos, por ejemplo, una sección de texto, una sección de tabla, una sección de imagen, etc., cada sección está sujeta a un conjunto diferente de reglas. El conjunto de reglas se usa para reparar los datos escaneados, generar segmentos hash y generar un hash de documento (etapas 3-4 en la página 4, líneas 12-16). A continuación, se recibe una firma digital (clave privada de un par de claves privada/pública en la etapa 6 en la página 4, líneas 19-20). Para fines de autenticación, se genera un código de autenticación y se imprime en el documento. Este código de autenticación contiene la firma digital, los hashes de segmento y el hash de documento, las reglas de reparación para cada uno de los segmentos y una forma digital de cada segmento (etapa 7 en la página 4, líneas 21-22). La forma digital también se genera en la etapa 3 al aplicar

el conjunto de reglas. Este conjunto de reglas depende del tipo de segmento, por ejemplo, "Texto", "Membrete", "logotipo", "Diagrama", "Tabla" o "Firma". Estas reglas intentan reparar los datos escaneados (consulte la página 6, líneas 7-36) para generar la forma digital y el hash.

El proceso de verificación de IBM requiere escanear el documento en papel, extraer los hashes de segmento y el hash de documento del documento escaneado y comparar estos hashes con los hashes de segmento y un hash de documento generado a partir del documento escaneado. Para evitar errores a través del escaneo, el método de IBM prevé también en el proceso de verificación una etapa intermedia en donde los datos escaneados se reparan, utilizando el mismo conjunto de reglas. La reparación la realiza manualmente el usuario que corrige los errores ortográficos o de contexto, o puede hacerse de forma semiautomática ignorando los avances de línea, reemplazando una serie de espacios con un solo espacio, etc. (vea la página 6, líneas 12-18 del documento EP 0 676 877). Los datos reparados se autentican con una mejor oportunidad para una autenticación exitosa en el caso de un documento válido y luego con el método de DeYoung.

Como resultado del escaneo y la segmentación, el proceso de autenticación conocido del documento EP 0 676 877 no se puede automatizar y se introducen errores en la primera etapa. El escaneo es una etapa manual que introduce errores de escaneo. Además, la segmentación es una etapa manual que generalmente se realiza con un editor y un ratón convencionales (consulte la página 4, líneas 9-12). Aunque no se excluye la posibilidad de reconocimiento automático de segmentos (consulte la página 5, líneas 44-47), esta etapa en general la realiza el usuario y también evita que se conozca el proceso del documento EP 0 676 877 de ser completamente automatizado. Además, la segmentación introduce más riesgos de errores. La segmentación es una etapa bastante compleja que requiere que la autoridad de firma seleccione datos de un solo tipo y asocie un conjunto de reglas con un editor convencional y un dispositivo de entrada. Aunque el método prevé una solución para hacer frente manual o semiautomáticamente a los errores de escaneo mediante la aplicación de un conjunto de reglas de reparación a los segmentos, el método sigue siendo propenso a errores.

El método de IBM permite autenticar solo documentos en papel y se ocupa de los problemas específicos que resultan de los documentos en papel. El método de IBM comienza con el papel (110 en la Figura 1) y produce papel (165 en la Figura 1). No hay enseñanza en el documento EP 0 676 877 sobre cómo se debe adaptar el método de IBM para autenticar eficientemente los documentos digitales originales. Una manera sencilla de autenticar un documento digital original sería imprimir una versión en papel del documento digital y alimentar la versión en papel al aparato 100 descrito por IBM en el documento EP 0,676,877, dejando al experto con un proceso que no se puede automatizar. Otra forma directa de autenticar un documento digital original sería quitar el escáner 125 del aparato 100 de IBM, dejando al experto con un proceso complejo que implica segmentación manual o semimanual, y códigos largos que incluyen conjuntos de reglas.

Es un objetivo de la presente invención describir un dispositivo y un proceso para generar documentos y un proceso correspondiente para verificar la autenticidad de los documentos, que resuelve las deficiencias descritas anteriormente de las soluciones de la técnica anterior. Más particularmente, es un objetivo presentar un dispositivo y proceso que permita autenticar eficientemente copias digitales y analógicas de documentos digitales originales sin tener que reproducir un hash a partir de un documento escaneado.

Resumen de la invención

De acuerdo con la presente invención, los objetivos identificados anteriormente son realizados por el dispositivo para proteger un documento digital original definido por la reivindicación 1, el dispositivo comprende:

- medios de recepción de documentos para recibir un documento digital original;
- medios de recepción de datos asociados para recibir datos asociados que comprenden al menos parcialmente datos visibles en el documento digital original;
- medios para crear una función hash para crear un hash de los datos asociados;
- medios de firma digital para cifrar asimétricamente el hash con una clave privada para generar datos firmados digitalmente;
- medios de codificación para generar un código legible por máquina que contiene el hash cifrado asimétricamente para firmar al menos una parte de los datos en el documento digital original, estando adaptados los medios de codificación para codificar en el código legible por máquina también los datos asociados de manera que el hash puede reproducirse a partir de datos asociados decodificados en un código legible por máquina de una copia impresa del documento digital original y servir para comparar con el hash cifrado codificado en el código legible por máquina de la copia impresa para verificar la autenticidad de los datos asociados decodificados, y el código legible por máquina permite la verificación de la autenticidad de la copia impresa del documento digital original mediante la comparación de datos en la copia impresa con los datos asociados decodificados del código legible por máquina y verificados; y
- medios de grabación para grabar el código legible por máquina en una versión digital protegida del documento digital original.

De hecho, al codificar en el código legible por máquina, tanto los datos asociados como un hash cifrado de los datos asociados, la verificación de la autenticidad ya no requiere la reproducción del hash de la copia impresa escaneada. Dado que los datos asociados en sí son parte del código legible por máquina, el hash puede reproducirse a partir de los datos asociados decodificados y servir para comparar con el hash cifrado codificado en el código legible por máquina. Además,

los datos relevantes en la copia impresa pueden compararse con los datos verificados asociados decodificados del código legible por máquina. La eficiencia y el rendimiento mejoran porque no todos los datos del documento están encriptados. Se obtiene una firma digital cifrando asimétricamente solo el hash con una clave privada.

Además de un dispositivo para proteger un documento digital original como se define en la reivindicación 1, la presente invención se refiere a un proceso correspondiente para proteger un documento digital original como se define en la reivindicación 8, el proceso comprende las siguientes etapas:

- recibir un documento digital original;
- recibir datos asociados que comprenden al menos parcialmente datos visibles en el documento digital original;
- crear un hash de los datos asociados;
- cifrar asimétricamente el hash con una clave privada para generar datos firmados digitalmente;
- generar un código legible por máquina que contenga el hash cifrado asimétricamente para firmar al menos una parte de los datos en el documento digital, y que contenga también los datos asociados de manera que el hash pueda reproducirse a partir de datos decodificados asociados en un código legible por máquina de un copia impresa del documento digital original y sirve para comparar con el hash cifrado codificado en el código legible por máquina de la copia impresa para verificar la autenticidad de los datos decodificados asociados, y el código legible por máquina permite verificar la autenticidad de la copia impresa del documento digital original mediante la comparación de datos en la copia impresa con los datos asociados decodificados a partir del código legible por máquina y verificados.
- grabar el código legible por máquina en una versión digital protegida del documento digital original;

La presente invención también se refiere a un proceso correspondiente para verificar la autenticidad de una copia impresa de un documento digital original como se define en la reivindicación 9, el proceso comprende:

- recibir una copia impresa del documento digital original que contiene un código legible por máquina en donde se codifica un hash cifrado asimétricamente de los datos asociados del documento digital original y los datos asociados en sí, los datos asociados comprenden al menos parcialmente datos visibles en el documento digital original;
- leer por máquina la copia impresa;
- decodificar los datos asociados del código legible por máquina;
- verificar si los datos asociados son auténticos reproduciendo el hash de los datos asociados decodificados a partir del código legible por máquina y comparando el hash reproducido con el hash cifrado codificado en el código legible por máquina, y para así obtener datos asociados verificados decodificados del código legible por máquina en caso de una coincidencia entre el hash reproducido y el hash cifrado codificado en el código legible por máquina; y
- comparar los datos asociados verificados decodificados del código legible por máquina con los datos extraídos de la copia impresa del documento para autenticar así la copia impresa en caso de una coincidencia entre los datos asociados verificados decodificados del código legible por máquina y los datos extraídos de la copia impresa.

Por lo tanto, verificar la autenticidad de una copia impresa del documento digital en su forma más simple consiste en comparar los datos relevantes en la copia impresa con los datos asociados verificados que se decodifican a partir del código legible por máquina en la versión protegida del documento digital.

Para verificar que los datos asociados decodificados del código legible por máquina sean auténticos, el hash de los datos asociados debe reproducirse a partir de los datos asociados decodificados y este hash reproducido debe compararse con el hash decodificado del código legible por máquina después de ser descifrado con la clave pública vinculada a la clave privada del firmante. Esto se define en la reivindicación 10.

De acuerdo con un aspecto innovador adicional de la invención, definido por la reivindicación 2, los medios de codificación en el dispositivo para proteger un documento digital original de acuerdo con la invención actual están además adaptados para codificar en el código legible por máquina, también una referencia a una clave pública o certificado asociado.

De hecho, en el proceso de verificación, el hash de los datos asociados que están codificados en la representación legible por máquina debe decodificarse y descifrarse utilizando la clave pública que corresponde con la clave privada utilizada en el proceso de protección de documentos. Esta clave pública se obtendrá preferentemente a través de una referencia codificada en el código legible por máquina.

Como se indica en la reivindicación 3, la representación legible por máquina en diferentes modalidades del dispositivo y proceso de acuerdo con la presente invención puede ser un código de barras, un código de barras 2D, un código holográfico, una banda magnética, un código de glifo de datos, un código de microdot, un código serpentino, una marca de agua, una etiqueta RFID, una URL, una secuencia alfanumérica legible usando OCR, un código de tinta magnética o una combinación de los anteriores. El código puede estar coloreado para aumentar la densidad de datos. Alternativamente, se puede utilizar un código invisible para humanos.

Además, opcionalmente, como se especifica en la reivindicación 4, los datos extraídos del documento digital original para la integración en el código legible por máquina pueden ser, por ejemplo, datos proporcionados junto con la representación digital del documento, pueden ser datos extraídos de una representación digital del documento al buscar marcadores, se

pueden extraer datos de una representación digital del documento al observar las posiciones predefinidas en el documento digital original, o una combinación de lo anterior.

Como se indica en la reivindicación 5 y la reivindicación 6 respectivamente, el dispositivo para proteger un documento digital de acuerdo con la presente invención puede integrarse en una aplicación de software de generación de documentos como Microsoft Word, o puede integrarse en un controlador de impresora.

Además opcionalmente, como se indica en la reivindicación 7, el dispositivo y el proceso de acuerdo con la invención pueden comprender uno o más de los siguientes:

- medios para codificar en el código legible por máquina una marca de tiempo;
- medios para codificar en el código legible por máquina una referencia indicativa de cómo se deben presentar los datos asociados a un usuario;
- medios para codificar en el código legible por máquina datos que no están presentes en forma legible por humanos en el documento digital original;
- medios para codificar en el código legible por máquina datos no firmados, como una referencia al documento digital original;
- medios para codificar en el código legible por máquina un nivel requerido de coincidencia antes de que una copia impresa se considere auténtica;
- medios para codificar en el código legible por máquina información indicativa de operaciones que se han aplicado a los datos asociados, como la compresión y codificación necesarias para verificar los datos asociados;
- medios para codificar en el código legible por máquina la firma digital del documento digital original.

Como se indica en la reivindicación 11, el proceso de verificación de acuerdo con la presente invención puede hacer uso de un escáner de escritorio, una cámara fotográfica, un escáner de código de barras o cualquier solución alternativa de hardware y/o software de lectura de documentos.

En el proceso para verificar la autenticidad de una copia impresa de acuerdo con la presente invención, las diferencias entre los datos asociados decodificados del código legible por máquina y los datos extraídos de la copia impresa pueden mostrarse opcionalmente al usuario. Esto se especifica en la reivindicación 12.

También opcionalmente, se puede mostrar al usuario un nivel de corroboración entre los datos asociados decodificados a partir del código legible por máquina y los datos extraídos de la copia impresa. Este aspecto opcional se define en la reivindicación 13.

En una posible implementación del proceso para verificar la autenticidad de una copia impresa de acuerdo con la presente invención, un operador puede estar involucrado como se indica en la reivindicación 14. Las preguntas relacionadas con el documento digital pueden hacerse a un operador y las respuestas recibidas pueden verificarse contra los datos asociados decodificados y verificados a partir de dicho código legible por máquina para validar la copia impresa.

Otra opción más del proceso de verificación de acuerdo con la presente invención implica generar una puntuación de autenticidad a partir de la comparación entre los datos asociados decodificados a partir del código legible por máquina y los datos extraídos de dicha copia impresa de dicho documento. El puntaje de autenticidad puede ser presentado al usuario. Este aspecto opcional se describe en la reivindicación 15.

## Breve descripción de los dibujos

La Figura 1 ilustra una modalidad del proceso y dispositivo para proteger un documento digital de acuerdo con la presente invención; y
la Figura 2 ilustra una modalidad del proceso correspondiente para verificar la autenticidad de una copia impresa del documento digital.

## Descripción detallada de las modalidades

La Figura 1 ilustra un proceso 10 que incorpora las etapas para proteger un documento digital 11. El documento digital original 11 se proporciona en la etapa 12 a un dispositivo, es decir, hardware o software que selecciona datos relevantes en el documento 11. Por lo general, se buscará en el documento texto relevante en posiciones específicas en el documento 11. En una modalidad más genérica, los datos asociados pueden recibirse junto con el documento original 11 como entrada al proceso. Los datos asociados en ese caso al menos incluyen una parte de los datos contenidos en el documento 11. En la etapa 13, se genera un hash a partir de los datos seleccionados o asociados a través de una función de hash especificada. En la etapa 15, este hash se firma digitalmente con una clave privada que se selecciona en la etapa 14. Por lo tanto, la firma digital implica un algoritmo de cifrado asimétrico basado en un par de claves privadas/públicas. La clave pública que se necesita en el proceso de verificación ilustrada en la Figura 2 puede tener un certificado asociado o de otra manera es confiable.

En la etapa 16, tanto los datos seleccionados como el hash firmado digitalmente se codifican en una representación legible por máquina. La representación legible por máquina, también llamada IntelliStamp™, puede ser, por ejemplo, un código de barras, un código de barras 2D, un código holográfico, una banda magnética, un código de glifo de datos, un código de microdot, un código de serpentina, una marca de agua, una etiqueta RFID, una URL, una secuencia alfanumérica legible usando OCR, un código de tinta magnética, etc., o una combinación de lo anterior. El código legible por máquina puede estar coloreado para aumentar la densidad de datos, o puede ser invisible para el usuario.

Los datos seleccionados y codificados en la representación legible por máquina pueden codificarse de manera que solo contengan caracteres ASCII legibles por humanos después de la decodificación. Se pueden colocar múltiples representaciones legibles por máquina en el documento para almacenar y codificar más información en los códigos. Además, los datos pueden registrarse de forma redundante en los códigos legibles por máquina. Por lo tanto, se pueden colocar múltiples representaciones legibles por máquina que contengan los mismos datos o datos parcialmente superpuestos en el documento para mejorar las posibilidades de recuperación de datos de un documento dañado.

En la etapa 17, la representación legible por máquina se registra en el documento original. La grabación o integración del código legible por máquina puede ser realizada por un sistema específico, o por un sistema integrado en el software de generación de documentos o software de controlador de impresora que genera una versión digital del documento que contiene el código legible por máquina.

Cualquier copia impresa del documento que se genera en la etapa 18 contendrá el código legible por máquina de manera que la autenticidad de la copia impresa pueda verificarse a través de un dispositivo o proceso como se ilustra en la Figura 2.

La Figura 2 ilustra un método 20 que incorpora etapas para validar la autenticidad de un documento impreso 21 generado por el método ilustrado en la Figura 1. En la etapa 22, la copia impresa se proporciona a un aparato que reconoce, captura y digitaliza representaciones legibles por máquina en el documento impreso para proporcionar información codificada. El aparato que lee el documento impreso en la etapa 22, más particularmente los códigos legibles por máquina, podría ser, por ejemplo, un escáner de escritorio, una cámara fotográfica, un escáner de código de barras, etc.

En la etapa 23, los datos asociados originales que se codificaron en los códigos legibles por máquina, se extraen de allí a través de la decodificación y los datos asociados decodificados se someten en la etapa 24 al algoritmo hash predefinido para reproducir el hash de los datos asociados decodificados.

En las etapas 25, 26 y 27, el hash de los datos asociados que se codificaron en la representación legible por máquina se decodifica desde allí y se descifra utilizando la clave pública que corresponde con la clave privada utilizada en el proceso ilustrado en la Figura 1. Esta clave pública se puede obtener, por ejemplo, en la etapa 26 directamente del firmante, de un tercero de confianza y/o puede venir con un certificado asociado.

El hash decodificado y descifrado de la representación legible por máquina se compara con el hash reproducido a partir de los datos asociados decodificados en la etapa 28. En caso de que no haya coincidencia, los datos asociados decodificables legibles por máquina no son auténticos como se indica en la etapa 30 y, en consecuencia, la copia impresa 21 no es auténtica. Cuando hay una coincidencia, los datos legibles por máquina decodificados son auténticos como se indica en la etapa 29.

En caso de que los datos asociados descifrados legibles por máquina decodificados sean auténticos, se compararán, ya sea automáticamente o por un operador, en la etapa 31 con los datos extraídos del propio documento impreso 21. La comparación llevará a la conclusión de que el documento impreso 21 es auténtico en la etapa 33 en caso de coincidencia, o llevará a la conclusión de que el documento impreso 21 no es auténtico en la etapa 32 en caso de que no haya coincidencia.

Opcionalmente, las diferencias entre los datos extraídos del documento impreso 21 y los datos decodificados legibles por máquina decodificados pueden mostrarse al operador, así como un nivel de corroboración entre ellos. Como resultado de la comparación, se puede calcular y compartir con el usuario un puntaje de autenticidad o un puntaje de autenticidad fraccional. La verificación en la etapa 31 puede involucrar a un operador al que se le hacen preguntas sobre el documento 21. Las respuestas del operador se pueden comparar con los datos asociados decodificados a partir de la representación legible por máquina. También se le puede pedir al operador que compare los datos asociados decodificados de la representación legible por máquina con la información visible en el documento impreso.

Aunque la presente invención se ha ilustrado como referencia a modalidades específicas, será evidente para los expertos en la técnica que la invención no se limita a los detalles de las modalidades ilustrativas anteriores, y que la presente invención puede llevarse a la práctica con varios cambios y modificaciones sin apartarse del alcance de la misma. Las modalidades presentes por lo tanto deben considerarse en todos los aspectos como ilustrativas y no restrictivas, el alcance de la invención se indica por las reivindicaciones adjuntas en vez de por la descripción anterior, y por lo tanto todos los cambios que entran dentro del ámbito de aplicación y la equivalencia de las reivindicaciones deben incluirse en el mismo. En otras palabras, se contempla cubrir todas y cada una de las modificaciones, variaciones o equivalentes que entren dentro del alcance de aplicación de los principios básicos subyacentes y cuyos atributos esenciales se reivindiquen en

esta solicitud de patente. Se deberá entender además por el lector de esta solicitud de patente que las palabras "que comprende" o "comprende" no excluyen otros elementos o etapas, que las palabras "uno" o "una" no excluyen una pluralidad, y que un único elemento, tal como un sistema informático, un procesador, u otra unidad integrada puede cumplir con las funciones de varios medios enumerados en las reivindicaciones. Cualquier signo de referencia en las reivindicaciones no deberá interpretarse como limitante de las reivindicaciones respectivas correspondientes. Los términos "primero", "segundo", "tercero", "a", "b", "c", y similares, cuando se usan en la descripción o en las reivindicaciones se introducen para distinguir entre elementos o etapas similares y no necesariamente describen un orden secuencial o cronológico. De manera similar, los términos "superior", "inferior", "sobre o a través", "bajo", y similares se introducen por propósitos descriptivos y no necesariamente para denotar posiciones relativas. Debe entenderse que los términos usados son intercambiables bajo circunstancias adecuadas y las modalidades de la invención son capaces de operar de acuerdo con la presente invención en otras secuencias, o en orientaciones diferentes de la(s) descrita(s) o ilustrada(s) anteriormente.

**REIVINDICACIONES**

1. Un dispositivo para proteger un documento digital original, dicho dispositivo comprende:
- medios de recepción de documentos para recibir (11) dicho documento digital original;
- medios de recepción de datos asociados para recibir (12) datos asociados que comprenden al menos parcialmente datos visibles en dicho documento digital original;
- medios para crear una función hash para crear (13) un hash de dichos datos asociados;
- medios de firma digital para cifrar asimétricamente (15) dicho hash con una clave privada para generar de este modo datos firmados digitalmente;
- medios de codificación para generar (16) un código legible por máquina que contiene dicho hash cifrado asimétricamente para firmar con al menos una porción de datos en dicho documento digital, dichos medios de codificación están adaptados para codificar en dicho código legible por máquina también dichos datos asociados de manera que dicho hash puede reproducirse a partir de datos decodificados asociados en un código legible por máquina de una copia impresa (21) de dicho documento digital original y servir para compararlo con el hash cifrado codificado en dicho código legible por máquina de dicha copia impresa (21) para verificar la autenticidad de dichos datos asociados decodificados, y dicho código legible por máquina permite la verificación de la autenticidad de dicha copia impresa (21) de dicho documento digital original mediante la comparación (31) de datos en dicha copia impresa con dichos datos asociados decodificados (23) de dicho código legible por máquina y verificados (24-30);
- medios de grabación para grabar (17) dicho código legible por máquina en una versión digital protegida de dicho documento digital original.

2. Un dispositivo para proteger un documento digital original de acuerdo con la reivindicación 1, en donde dichos medios de codificación están adaptados además para codificar en dicho código legible por máquina una referencia a una clave pública o certificado asociado.

3. Un dispositivo para proteger un documento digital original de acuerdo con la reivindicación 1, caracterizado porque dicho código legible por máquina es uno o más de:
- un código de barras;
- un código de barras 2D;
- un código holográfico;
- una banda magnética;
- un código de glifo de datos;
- un código de microdot;
- un código serpentino;
- una marca de agua;
- una etiqueta RFID;
- una URL;
- una secuencia alfanumérica legible usando OCR;
- un código de tinta magnética;
- un código de color;
- un código invisible para humanos.

4. Dispositivo para proteger un documento digital original de acuerdo con la reivindicación 1, caracterizado porque dichos datos extraídos de dicho documento digital original comprenden uno o más de:
- datos junto a dicho documento digital original;
- datos identificados por marcadores en dicho documento digital original;
- datos en una posición predefinida en dicho documento digital original.

5. Dispositivo para proteger un documento digital original de acuerdo con la reivindicación 1, caracterizado porque dicho dispositivo está integrado en una aplicación de software de generación de documentos.

6. Dispositivo para proteger un documento digital original de acuerdo con la reivindicación 1, caracterizado porque dicho dispositivo está integrado en un controlador de impresora.

7. Dispositivo para proteger un documento digital original de acuerdo con la reivindicación 1, caracterizado porque dicho dispositivo comprende además uno o más de los siguientes:
- medios para codificar en dicho código legible por máquina una marca de tiempo;
- medios para codificar en dicho código legible por máquina una referencia indicativa de cómo deben presentarse dichos datos asociados a un usuario;
- medios para codificar en dicho código de datos legibles por máquina no presentes en forma legible por humanos en dicho documento digital original;
- medios para codificar en dicho código legible por máquina datos no firmados tales como una referencia a dicho documento digital original;
- medios para codificar en dicho código legible por máquina un nivel requerido de coincidencia antes de que una copia impresa se considere auténtica;

- medios para codificar en dicha información de código legible por máquina indicativa de operaciones que se han aplicado a dichos datos asociados, como la compresión y codificación necesarias para verificar dichos datos asociados;
- medios para codificar en el código legible por máquina la firma digital del documento digital original.

8. Un proceso para proteger un documento digital original, dicho proceso comprende:
- recibir (11) dicho documento digital original;
- recibir (12) datos asociados que comprenden al menos parcialmente datos visibles en dicho documento digital original;
- crear (13) un hash de dichos datos asociados;
- cifrar asimétricamente (15) dicho hash con una clave privada para generar así datos firmados digitalmente;
- generar (16) un código legible por máquina que contenga dicho hash cifrado asimétricamente para firmar al menos una parte de los datos en dicho documento digital original y que contenga también dichos datos asociados de manera que dicho hash pueda reproducirse a partir de datos asociados decodificados en un código legible por máquina de una copia impresa (21) de dicho documento digital original y sirve para compararlo con el hash cifrado codificado en dicho código legible por máquina de dicha copia impresa (21) para verificar la autenticidad de dichos datos decodificados asociados, y dicha código legible por máquina de dicha copia impresa permite la verificación de la autenticidad de dicha copia impresa (21) de dicho documento digital original mediante la comparación de datos en dicha copia impresa (21) con dichos datos asociados decodificados a partir de dicho código legible por máquina y verificados; y
- grabar (17) dicho código legible por máquina en una versión digital protegida de dicho documento digital original.

9. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original,
dicho proceso comprende:
- recibir (21) una copia impresa de dicho documento digital original que contiene un código legible por máquina en donde se codifican un hash cifrado asimétricamente de datos asociados del documento digital original y dichos datos asociados, dichos datos asociados comprenden al menos parcialmente datos visibles en dicho documento digital original;
- leer por máquina (22) dicha copia impresa;
- decodificar (23) dichos datos asociados de dicho código legible por máquina; y
- verificar (24-30) si dichos datos asociados son auténticos al reproducir dicho hash de dichos datos asociados decodificados a partir de dicho código legible por máquina y al comparar dicho hash reproducido con dicho hash cifrado codificado en dicho código legible por máquina, y para obtener así datos asociados decodificados verificados a partir de dicho código legible por máquina en caso de una coincidencia entre dicho hash reproducido y dicho hash cifrado codificado en dicho código legible por máquina;
- comparar (31) dichos datos asociados verificados decodificados a partir de dicho código legible por máquina con los datos extraídos de dicha copia impresa de dicho documento para autenticar de ese modo dicha copia impresa en caso de una coincidencia entre dichos datos asociados verificados decodificados a partir de dicho código legible por máquina y dichos datos extraídos de dicha copia impresa.

10. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original de acuerdo con la reivindicación 9,
caracterizado porque dicha etapa de verificar (24-30) si dichos datos asociados son auténticos comprende:
- decodificar (25) dicho hash cifrado asimétricamente de dichos datos asociados a partir de dicho código legible por máquina para obtener de ese modo un hash decodificado cifrado asimétricamente;
- descifrar (26-27) dicho hash decodificado encriptado asimétricamente con una clave pública asociada con la clave privada de los firmantes para obtener así un hash descifrado;
- generar (24) un hash a partir de dichos datos asociados decodificados a partir de dicho código legible por máquina; y
- comparar (28) dicho hash descifrado con dicho hash generado a partir de dichos datos asociados decodificados de dicho código legible por máquina para verificar de ese modo la autenticidad de dichos datos asociados decodificados de dicho código legible por máquina.

11. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original de acuerdo con la reivindicación 9,
caracterizado porque leer por máquina dicha copia impresa se realiza utilizando uno de los siguientes:
- un escáner de escritorio;
- una cámara fotográfica;
- un escáner de código de barras.

12. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original de acuerdo con la reivindicación 9,
caracterizado porque dicho proceso comprende además:
- mostrar diferencias entre dichos datos asociados decodificados a partir de dicho código legible por máquina y los datos extraídos de dicha copia impresa de dicho documento digital original.

13. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original de acuerdo con la reivindicación 12,
caracterizado porque dicho proceso comprende además:
- determinar y mostrar un nivel de corroboración entre dichos datos asociados decodificados a partir de dicho código legible por máquina y los datos extraídos de dicha copia impresa de dicho documento digital original.

14. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original de acuerdo con la reivindicación 9,
caracterizado porque dicho proceso comprende además:
- hacer preguntas sobre dicho documento digital original a un operador;
- recibir respuestas a dichas preguntas de dicho operador; y
- verificar dichas respuestas contra dichos datos asociados decodificados a partir de dicho código legible por máquina.

15. Un proceso para verificar la autenticidad de una copia impresa de un documento digital original de acuerdo con la reivindicación 9,
caracterizado porque dicho proceso comprende además:
- generar una puntuación de autenticidad a partir de la comparación entre dichos datos asociados decodificados a partir de dicho código legible por máquina y dichos datos extraídos de dicha copia impresa de dicho documento digital original.

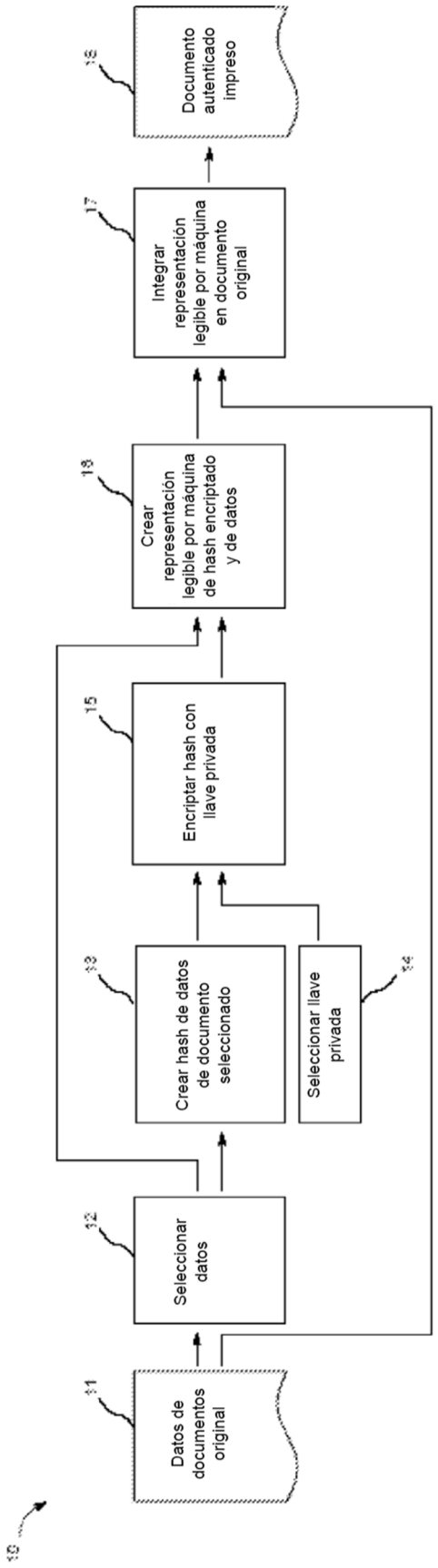
10
11
Datos de documentos original
12
Seleccionar datos
13
Crear hash de datos de documento seleccionado
14
Seleccionar llave privada
15
Encriptar hash con llave privada
16
Crear representación legible por maquina de hash encriptado y de datos
17
Integrar representación legible por maquina en documento original
18
Documento autenticado impreso

Fig. 1

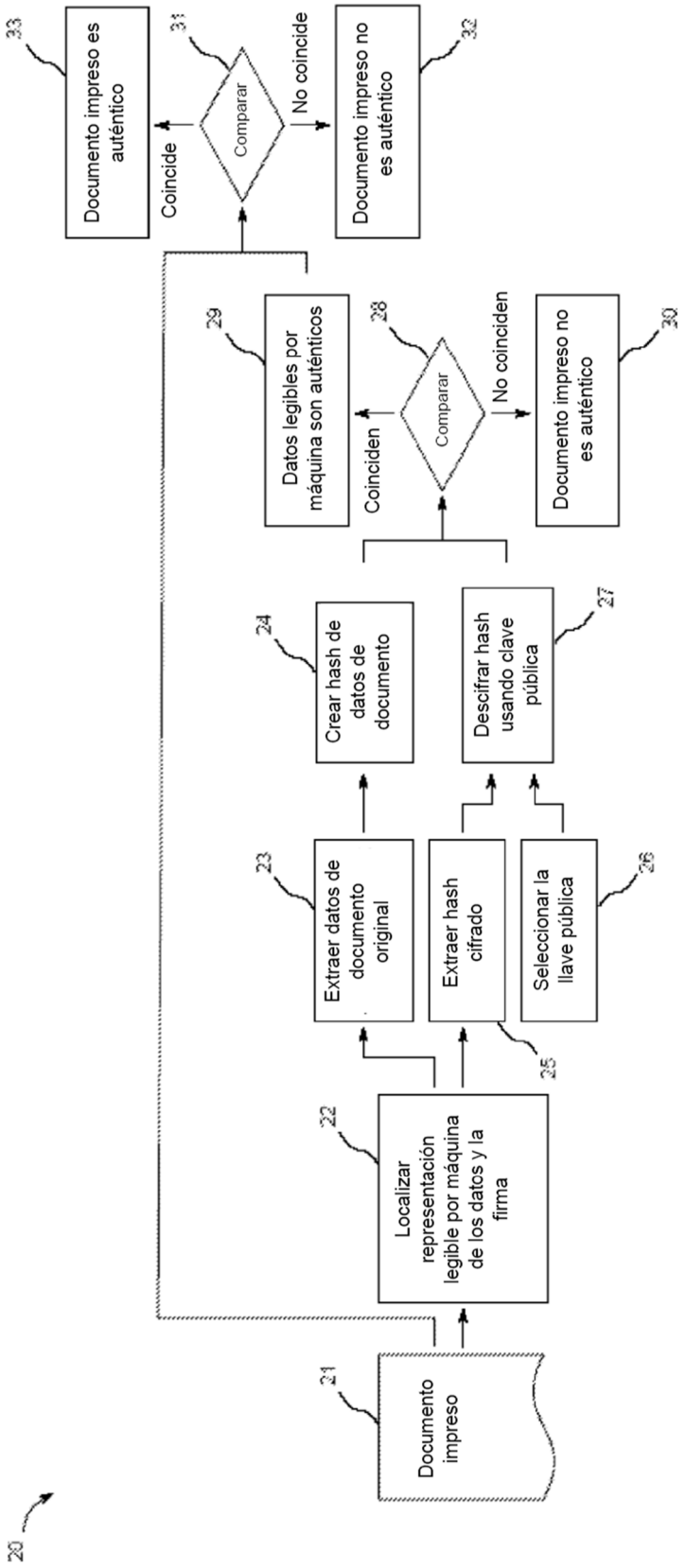
20
21
Documento impreso
22
Localizar representación legible por máquina de los datos y la firma
23
Extraer datos de documento original
24
Crear hash de datos de documento
25
Extraer hash cifrado
26
Seleccionar la llave pública
27
Descifrar hash usando clave pública
28
Comparar
Coinciden
No coinciden
29
Datos legibles por máquina son auténticos
30
Documento impreso no es auténtico
31
Comparar
Coincide
No coincide
32
Documento impreso no es auténtico
33
Documento impreso es auténtico

Fig. 2