

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 668 772**

51 Int. Cl.:

**G06F 17/30** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.10.2014 PCT/IB2014/065507**

87 Fecha y número de publicación internacional: **30.04.2015 WO15059637**

96 Fecha de presentación y número de la solicitud europea: **21.10.2014 E 14806721 (8)**

97 Fecha y número de publicación de la concesión europea: **28.02.2018 EP 3061007**

54 Título: **Método para la generación de una ontología informática, producto de programa informático configurado para implementar las operaciones de dicho método y equipo informático configurado para ejecutar dicho producto de programa informático**

30 Prioridad:

**22.10.2013 IT VI20130258**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:

**22.05.2018**

73 Titular/es:

**2C SOLUTION S.R.L. (100.0%)  
Via Longhin 1/5  
35011 Campodarsego (PD), IT**

72 Inventor/es:

**COLETTO, DAVIDE y  
PIAZZA, MIRKO**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 668 772 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

**DESCRIPCIÓN**

Método para la generación de una ontología informática, producto de programa informático configurado para implementar las operaciones de dicho método y equipo informático configurado para ejecutar dicho producto de programa informático

5 La invención versa sobre un método mejorado para la generación de una ontología informática mediante el análisis de varios documentos digitales con contenido alfanumérico almacenados en uno o más soportes de almacenamiento.

10 La invención también versa sobre un producto de programa informático configurado para implementar las operaciones de dicho método de la invención cuando dicho producto es ejecutado en un dispositivo electrónico de procesamiento.

Por último, la invención también versa sobre un aparato informático que comprende un dispositivo electrónico de procesamiento configurado para ejecutar dicho producto de programa informático y, así, implementar las operaciones de dicho método de la invención.

15 Es sabido que, en el campo de la informática, el término “ontología” significa un vocabulario estructurado que representa un conjunto de metadatos relacionados entre sí y relacionados con un dominio de aplicación específico.

El uso de una ontología, por lo tanto, permite a un usuario o a una aplicación informática obtener la información deseada mediante la formulación de “consultas” de alto nivel que son traducidas de inmediato, por medio de procesos lógicos especiales, en interrogaciones de bajo nivel a base de datos o a las bases de datos en las que está contenida dicha información.

20 Por ejemplo, dichas ontologías pueden ser usadas para reconocer automáticamente el contenido alfanumérico de un documento digital para poder catalogarlo por área temática. De esta manera, es posible transmitir dicho documento analizado a una ubicación de memoria o a un usuario preciso perteneciente a la misma área temática.

Sin embargo, también es sabido que, de forma desventajosa, las operaciones para crear dichas ontologías son muy complejas y precisan un esfuerzo considerable en términos de recursos y tiempo.

25 En particular, de forma desventajosa, en la actualidad, una ontología fiable que permita obtener resultados precisos sigue necesitando una aportación humana determinante nada insignificante.

En particular, la generación de una ontología requiere una aportación considerable por parte de personal experto que tiene que dedicar su pericia intelectual y su tiempo a la definición y a la asociación de los términos que definirán dicho vocabulario estructurado. Por lo tanto, el coste económico para la definición de una ontología es bastante alto.

30 En “A Concept-Driven Automatic Ontology Generation Approach for Conceptualization of Document Corpora”, IEEE Conference on Web Intelligence and Intelligent Agent Technology, diciembre de 2008, Hai-Tao Zheng *et al* dieron a conocer un método para derivar automáticamente una ontología a partir del contenido de documentos, independientemente de la organización del almacenamiento de los documentos. La presente invención se propone superar las desventajas mencionadas más arriba.

35 En particular, uno de los objetos de la presente invención es definir un método que hace posible generar automáticamente una ontología partiendo de documentos digitales con contenido alfanumérico almacenados en uno o más soportes de almacenamiento.

Por lo tanto, un objeto de la presente invención es definir un método automático para analizar dichos documentos digitales para la creación de una ontología que elimina la necesidad de aportaciones humanas.

40 En consecuencia, un objeto de la invención es definir un método que hace posible reducir el tiempo y los recursos necesarios para la creación de dicha ontología.

Otro objeto no menor de la invención es definir un método para la generación de una antología que hace posible catalogar un documento específico de manera más precisa en comparación con las ontologías generadas con los métodos conocidos.

45 Los objetos mencionados en lo que antecede se logran por medio del método para la creación de una antología llevado a cabo según la reivindicación principal.

50 En particular, el método de la invención para la creación automática de una antología se caracteriza porque asocia la información obtenida de la ruta de cada documento digital con la información obtenida del análisis del contenido del mismo documento. Esta asociación posibilita, de forma ventajosa, definir palabras clave (en jerga técnica también denominadas “identificadores”) que se considera que son de primer nivel y que corresponden a los nombres de las carpetas virtuales obtenidas de la ruta de cada documento digital, asociadas con palabras clave (“identificadores”) de segundo nivel que, a su vez, corresponden a porciones del contenido de dicho documento

digital. En particular, dichas palabras clave de primer nivel serán usadas como macrocategorías para la catalogación de nuevos documentos digitales por medio de dicha ontología, mientras que las palabras clave de segundo nivel serán usadas como subcategorías en la misma catalogación.

En las reivindicaciones dependientes se describen características adicionales del método de la invención.

5 La invención también incluye el producto de programa informático configurado para implementar las operaciones de dicho método de la invención y el equipo informático que comprende un dispositivo electrónico de procesamiento configurado para ejecutar dicho producto de programa informático, cuyas características son descritas, respectivamente, en las reivindicaciones 11 y 12.

10 Dichos objetos, junto con las ventajas que serán destacadas posteriormente, son ilustrados con mayor detalle en la descripción de una realización preferente de la invención que es proporcionada a título de ejemplo no limitante con referencia a los dibujos adjuntos, en los que:

- la Figura 1 muestra el diagrama de bloques del método de la invención para generar automáticamente una ontología informática.

15 El método que es el objeto de la invención para la generación de una ontología informática mediante el análisis de varios documentos digitales con contenido alfanumérico está representado en conjunto con el diagrama de bloques mostrado en la Figura 1, en la que está indicado por **1**.

20 Dichos documentos digitales están almacenados en uno o más soportes de almacenamiento, como, por ejemplo, discos duros integrados en un dispositivo de procesamiento configurado para implementar el método **1** de la invención, o, alternativamente, discos duros externos conectados a dicho dispositivo electrónico de procesamiento, pudiendo estar ubicados dichos discos de forma tanto local como remota ("nube").

Además, o como alternativa de lo anterior, dichos documentos digitales pueden estar almacenados en diferentes tipos de soportes de almacenamiento de tipo portátil conectados temporalmente a dicho dispositivo electrónico de procesamiento, como, por ejemplo, los conocidos lápices USB.

25 En cuanto al tipo de documentos digitales que han de ser analizados, estos pueden ser documentos digitales de texto, de modo que el contenido alfanumérico pueda ser abordado directamente por medio del método **1** de la invención. Alternativamente, dichos documentos pueden estar en formato de imagen digital, de modo que, antes de poder analizar su contenido alfanumérico, sea necesario llevar a cabo una operación intermedia, que suele incluir la aplicación de un algoritmo de tipo OCR (reconocimiento óptico de caracteres).

30 Según puede observarse en la Figura 1, el método **1** de la invención incluye, para cada uno de los documentos digitales almacenados en dichos soportes de almacenamiento, la operación **2** de extracción de su contenido alfanumérico y la operación **3** de análisis de este, de tal modo que se identifique cada uno de sus lexemas.

En esta descripción, el término "lexema" significa cada conjunto de caracteres indivisibles.

35 En particular, en lo referente a dicha operación **3** de análisis del contenido alfanumérico, en el campo de la informática esto es conocido como "análisis léxico". Con mayor detalle, esta operación incluye la definición y la identificación, dentro de dicho contenido alfanumérico, de uno o más elementos separadores que distinguen entre un lexema específico y el adyacente. Según la realización preferente de la invención descrita en la presente memoria, dichos elementos separadores son los espacios vacíos entre caracteres y los signos de puntuación. Sin embargo, no puede descartarse que, en diferentes realizaciones del método **1** de la invención, puedan seleccionarse diferentes elementos separadores, siempre y cuando hagan posible distinguir entre un lexema y los adyacentes.

40 Una vez que cada lexema individual ha sido identificado dentro de un documento digital específico según el método **1**, que es el objeto de la invención, cada uno de dichos lexemas es concatenado con uno o más lexemas adyacentes a él, de tal modo que se definan varias frases. Dicha operación está indicada por **4** en la Figura 1.

45 Según el método **1** de la invención, el mínimo número y el máximo número de lexemas que pueden constituir frases están definidos de antemano. Según la realización preferente de la invención, se ha establecido que dicho número mínimo de lexemas sea dos y que dicho número máximo de lexemas sea cinco. Obviamente, en realizaciones diferentes del método **1** de la invención, tanto el número mínimo como el número máximo de lexemas pueden ser inicializados con valores diferentes de los adoptados en la realización preferente de la invención descrita en la presente memoria.

50 Además, según dicha realización preferente de la invención, las frases que tienen, al comienzo o al final, uno o más términos definidos de antemano como inadmisibles son excluidas de las frases definidos mediante la operación de concatenación anteriormente mencionada. Un ejemplo de términos inadmisibles puede estar representado por el grupo de artículos definidos e indefinidos constituido por los lexemas "el", "la", "los", "las", "un" y "una".

Realizaciones alternativas variantes del método **1** de la invención pueden no permitir tal exclusión, o puede permitir la exclusión de las frases que contengan términos que sean considerados inadmisibles.

Según el método **1** de la invención, una vez que se ha completado la operación **4** de definición de las frases, incluida la etapa de exclusión, se calcula el número de veces que aparece cada una de dichas frases dentro del documento digital específico. Dicha operación es indicada por **5** en la Figura 1.

5 En otras palabras, el método calcula cuántas veces se repite la misma frase dentro del documento digital que se analiza.

Dicho cálculo de las veces que aparecen las frases hace posible llevar a cabo la operación subsiguiente de clasificar las mismas frases, indicada por **6** en la Figura 1. De hecho, según el método **1** de la invención, las frases son clasificadas en orden descendente en función del número de apariciones.

10 En particular, preferente pero no necesariamente, según la realización preferente de la invención se calcula un índice de clasificación  $i$  para cada frase y corresponde, de hecho, al número de veces que aparece cada frase dividido por el número total de lexemas presentes en el documento digital y multiplicado por el número de lexemas que componen la frase en cuestión. Por lo tanto, en este caso las frases son clasificadas en orden descendente en función de dicho índice de clasificación  $i$ .

15 El resultado de dicha operación de clasificación es almacenado en una primera ubicación **100** de memoria en uno de dichos soportes de almacenamiento. Dicha operación de almacenamiento es indicada por **7** en la Figura 1. Según la invención, el método **1** para la generación de una ontología informática incluye, además, la operación **8** of extracción de la cadena relativa a la ruta del documento digital. Esta cadena es analizada subsiguientemente, y esta operación es indicada por **9** en la Figura 1, de tal modo que se identifique el nombre identificativo de cada carpeta virtual individual de dicha ruta. En aras de la claridad, de este punto de la descripción en adelante, el término "carpeta" o "carpetas" por sí solo será usado para referirse a carpetas virtuales.

20 Preferente, pero no necesariamente, según la realización preferente del método **1** de la invención, dicha operación de análisis de dicha cadena considera los signos "/" o "\" (también denominados, respectivamente, "barra" y "barra inversa") como elementos separadores entre cada uno de los nombres de carpeta y los adyacentes. Obviamente, no puede descartarse que en diferentes realizaciones del método **1** de la invención el elemento de separación seleccionado pueda ser diferente de los recién indicados con referencia a la primera realización de la invención.

Habitualmente, la selección de dichos elementos separadores depende del sistema operativo y del "sistema de ficheros" presente en el equipo informático en el que se implemente el método **1** de la invención y en el que se almacenen dichos documentos digitales.

30 Según el método **1** de la invención, una vez que dicha operación **9** de identificación de los nombres de carpeta presentes en la ruta de dicho documento digital que se está analizando, cada uno de dichos nombres de carpeta es indexado. Dicha operación es indicada por **10** en la Figura 1.

35 En particular, según el método **1** de la invención, la operación **10** incluye, en primer lugar, la selección del nombre de carpeta a partir del cual tiene que empezar la operación de indexado. Preferente, pero no necesariamente, la selección del nombre de carpeta para el inicio de dicha operación de indexado debería comenzar a partir del nombre de la carpeta de la rama más exterior, lo que significa la carpeta a la que está realmente asignado el documento digital considerado, retrocediendo jerárquicamente en la ruta un número  $n$  de nombres de carpeta definido de antemano. Por ejemplo, si la ruta de un documento digital específico comprende una estructura jerárquica que consiste en 10 nombres de carpeta y dicho número  $n$  seleccionado es 2, se selecciona el octavo nombre de carpeta de dicha ruta ( $10-2 = 8$ ) como nombre de carpeta inicial para la operación de indexado. En este punto, la operación de indexado **10** requiere el indexado de dicho nombre de carpeta inicial y de los nombres relacionados con dichas subcarpetas con un valor creciente de 1 a  $n$ , empezando a partir de dicho nombre de carpeta inicial y prosiguiendo hacia las subcarpetas relacionadas según el orden jerárquico de dicha ruta. Según se ha explicado anteriormente, si dicho valor  $n$  seleccionado es dos, el nombre de carpeta inicial es indexado con el valor 1, mientras que la subcarpeta adyacente a la carpeta inicial es indexada con el valor 2.

45 Obviamente, a  $n$  se le puede asignar un valor numérico entero mayor que 2. En este caso, se indexarán más subcarpetas en comparación con las indexadas según el ejemplo propuesto anteriormente.

Según el método **1** de la invención, una vez que se ha completado la operación **10** de indexado, cada uno de los nombres de carpeta con valor  $m$  es asociado, mediante la operación indicada por **11** en la Figura 1, con el correspondiente nombre de carpeta con el valor  $m+1$ , estando el valor  $m$  incluido entre 1 y  $n-1$  ( $1 \leq m \leq n-1$ ).

50 El método **1** de la invención, en particular, comienza a partir del nombre de carpeta con valor 1, correspondiente al nombre de carpeta inicial mencionado anteriormente, y lo asocia con el nombre de carpeta con valor 2, y luego, considera, a su vez, que el nombre de carpeta con valor 2 está asociado con el nombre de carpeta subsiguiente, etcétera, aumentando progresivamente  $m$  en una unidad hasta obtener  $m-n-1$ .

55 Según el método **1** de la invención, una vez que se ha completado dicha operación **11** de asociación, se lleva a cabo la operación **12**, que es una operación de almacenamiento de los resultados de dichas operaciones de indexado y

asociación en una segunda ubicación **200** de memoria, compartida entre la totalidad de dichos documentos digitales analizados.

5 De esta manera, analizando progresivamente los diversos documentos digitales, es posible definir uno o más sinónimos constituidos por los nombres de carpeta con valor  $m+1$  para cada uno de dichos nombres de carpeta con valor  $m$ , estando asociados dichos nombres con valor  $m+1$ , mediante la precedente operación **11** de asociación, con el nombre de carpeta anteriormente mencionado con valor  $m$ . De nuevo, se prevé que el valor  $m$  esté incluido entre 1 y  $n-1$  ( $1 \leq m \leq n-1$ ).

10 En otras palabras, se obtiene una estructura jerárquica, según la cual se define un conjunto de nombres relacionados con subcarpetas con valor  $m+1$  para cada nombre de carpeta con valor  $m$ , considerándose dichos nombres de subcarpetas sinónimos mutuos en el momento en que la ontología generada de esta manera es usada para para la catalogación de los nuevos documentos digitales.

15 Por último, el método **1** de la invención incluye la operación adicional **13** de asociación de los nombres de carpeta almacenados en dicha segunda ubicación **200** de memoria, incluyendo los sinónimos relacionados y los valores de indexado, con un número predefinido  $x$  de dichas frases clasificadas, a su vez almacenadas en la primera ubicación **100** de memoria.

En particular, esta operación de asociación **13** comienza a partir de la primera frase de dicha lista de frases clasificadas que corresponde, según la realización preferente de la invención, a la frase con el mayor índice de clasificación  $i$  entre todas las frases presentes en el documento digital que se analiza. Sucesivamente, las frases presentes en dicha lista son consideradas, de manera ordenada y progresiva, hasta alcanzar la sentencia  $x$ -ésima.

20 Según la realización preferente del método **1** de la invención, el número de frases  $x$  es seleccionado en función del número total de lexemas contenidos en el documento digital que se analiza. En particular, preferente, pero no necesariamente, el número de frases  $x$  corresponde a 15 frases cuando el número total de lexemas contenidos en el documento es mayor que 1000, o dicho número de frases  $x$  corresponde a 10 frases cuando dicho número total de lexemas está incluido entre 1000 y 101. Por último, si el documento contiene un número de lexemas igual o menor que 100, el número de frases  $x$  seleccionado es 5.

Además, en cuanto a dicha operación **13** de asociación, los nombres de carpeta que están asociados con cada frase corresponden a los nombres de carpeta que pueden hallarse realmente en la ruta del documento digital que se analiza.

30 Al final de dicha operación **13**, se obtiene de forma ventajosa la ontología deseada mediante un procedimiento automático, logrando así los objetos de la invención. Esta ontología puede ser usada entonces para la catalogación de los nuevos documentos digitales cargados en el equipo informático en cuestión.

En particular, para poder poner dicha ontología a la disposición de aplicaciones informáticas o de diversos usuarios, el resultado de dicha operación **13** de asociación es almacenado en una tercera ubicación **300** de memoria, accesible al público general o mediante credenciales de acceso. Esta operación es indicada por **14** en la Figura 1.

35 La invención también incluye el producto de programa informático que comprende una porción de código de soporte lógico directamente cargable en la memoria de un dispositivo electrónico de procesamiento y configurada para implementar las operaciones de dicho método **1** de la invención cuando dicha porción de código de soporte lógico es ejecutada en dicho dispositivo electrónico de procesamiento.

40 Además, la invención también incluye el equipo informático que comprende un dispositivo electrónico de procesamiento dotado de uno o más soportes de almacenamiento en los cuales están almacenados uno o más de dichos documentos digitales. Según la invención, también dicha porción de código de soporte lógico del producto de programa informático de la invención está almacenada en dichos soportes de almacenamiento.

El dispositivo electrónico de procesamiento comprende, además, un procesador configurado para ejecutar dicha porción de código de soporte lógico de tal modo que implemente las operaciones del método **1** de la invención.

45 Por lo tanto, según la explicación proporcionada en lo que antecede, el método **1**, el producto de programa informático y el equipo informático de la invención logran todo el conjunto de objetos.

En particular, la invención logra el objeto de definir un método que hace posible generar automáticamente una ontología partiendo de documentos digitales con contenido alfanumérico almacenados en uno o más soportes de almacenamiento.

50 Por lo tanto, la invención también logra el objeto de definir un método automático para el análisis de dichos documentos digitales para crear una ontología que elimine la necesidad de aportaciones humanas.

En consecuencia, un objeto adicional logrado por la invención es el objeto de definir un método que hace posible reducir el tiempo y los recursos necesarios para la creación de dicha ontología.

Por último, la invención también logra el objeto de definir un método para la generación de una antología que hace posible catalogar un documento específico de manera más precisa en comparación con las ontologías generadas con los métodos conocidos.

## REIVINDICACIONES

1. Un método (1) implementado por ordenador para la generación de una ontología informática mediante el análisis de varios documentos digitales de contenido alfanumérico almacenados en uno más soportes, incluyendo dicho método (1) la realización de las siguientes operaciones para cada uno de dichos documentos:

5 - extraer dicho contenido alfanumérico de dicho documento digital y analizar dicho contenido de tal modo que se identifique cada uno de sus lexemas;

- concatenar cada uno de dichos lexemas con uno o más lexemas adyacentes, de tal modo que se definan varias frases, estando definido de antemano el número mínimo y máximo de lexemas que pueden estar contenidos en cada frase;

10 - calcular el número de veces que aparece cada una de dichas frases en dicho documento digital;

- clasificar dichas frases en orden decreciente según dichas apariciones calculadas y almacenar dichas frases clasificadas en una primera ubicación (100) de memoria;

**caracterizado porque** incluye la realizaciones de las siguientes operaciones adicionales para cada uno de dichos documentos:

15 - extraer la cadena relacionada con la ruta de dicho documento digital, identificando dicha ruta una ubicación en la que está almacenado dicho documento digital en el uno o más soportes de almacenamiento, y analizar dicha cadena de tal modo que se identifique el nombre identificativo de cada carpeta individual de cada ruta;

- seleccionar, entre dichos nombres de carpeta identificados, un nombre de carpeta inicial a partir del cual sea posible proseguir con la siguiente operación de indexado;

20 - indexar dichos nombres de carpeta, comenzando con dicho nombre de carpeta inicial y continuando hacia las correspondientes subcarpetas según el orden jerárquico de dicha ruta, aumentando el valor numérico entero de 1 a  $n$ ;

- asociar cada uno de dichos nombres de carpeta con valor  $m$  con el nombre de carpeta relacionado con el valor  $m+1$ , siendo  $1 \leq m \leq n-1$  y aumentando  $m$  progresivamente en una unidad;

25 - almacenar los resultados de dichas operaciones de indexado y asociación en una segunda ubicación (200) de memoria compartida por todos dichos documentos digitales analizados, de tal modo que se definan progresivamente, para cada uno de dichos nombres de carpeta con valor  $m$ , uno o más sinónimos constituidos por los nombres de carpeta con valor  $m+1$  asociados, mediante la operación anterior, con dicho nombre de carpeta con valor  $m$ , siendo  $1 \leq m \leq n-1$ ;

30 - asociar dichos nombres de carpeta almacenados en dicha segunda ubicación (200) de memoria, incluyendo los correspondientes sinónimos y los valores de indexado, con un número predefinido  $x$  de dichas frases clasificadas, comenzando con la primera de dichas frases y continuando de manera ordenada y progresiva, correspondiendo dichos nombres de carpeta asociados a los nombres de carpeta presentes en dicha ruta de dicho documento digital;

- almacenar el resultado de dichas operaciones de asociación en una tercera ubicación (300) de memoria.

35 **2.** Un método (1) según la reivindicación 1 **caracterizado porque**, entre dichas varias frases definidas durante dicha operación de concatenación, se excluyen las frases que tienen, al comienzo o al final, uno o más términos definidos de antemano como inadmisibles.

40 **3.** Un método (1) según cualquiera de las reivindicaciones precedentes **caracterizado porque** dicha operación de clasificación de dichas frases incluye el cálculo, para cada una de dichas frases, de un índice de clasificación  $i$  correspondiente al número de veces que aparece dicha frase en dicho documento digital, dividido por el número total de lexemas de dicho documento y multiplicado por el número de lexemas que componen dicha frase, clasificándose dichas frases en orden decreciente en función de dicho índice de clasificación  $i$ .

45 **4.** Un método (1) según cualquiera de las reivindicaciones precedentes **caracterizado porque** dicho número mínimo de lexemas que cada frase puede contener es dos, y dicho número máximo de lexemas que cada frase puede contener es cinco.

**5.** Un método (1) según cualquiera de las reivindicaciones precedentes **caracterizado porque** dicho número predefinido  $n$  es igual a dos.

**6.** Un método (1) según cualquiera de las reivindicaciones precedentes **caracterizado porque** dicho número de frases  $x$  es seleccionado en función del número total de lexemas contenidos en dicho documento digital.

50 **7.** Un método (1) según la reivindicación 6 **caracterizado porque** dicho número de frases  $x$  corresponde a 15 frases cuando dicho número total de lexemas es mayor de 1000, o dicho número de frases  $x$  corresponde a 10 frases cuando dicho número total de lexemas está incluido entre 1000 y 101, o dicho número de frases  $x$  corresponde a 5 frases cuando dicho número total de lexemas es igual o menor que 100.

55 **8.** Un método (1) según cualquiera de las reivindicaciones precedentes **caracterizado porque**, según dicha operación de análisis de dicho contenido de dicho documento digital, adecuada para identificar cada lexema

individual, los espacios vacíos y los signos de puntuación son considerados elementos separadores entre cada uno de dichos lexemas y los adyacentes.

5 **9.** Un método (1) según cualquiera de las reivindicaciones precedentes **caracterizado porque**, según dicha operación de análisis de dicha cadena relativa a dicha ruta, adecuada para identificar el nombre identificativo de cada carpeta individual, los símbolos "/" o "\" son considerados elementos separadores entre cada uno de dichos nombres de carpeta y los adyacentes.

10 **10.** Un método según cualquiera de las reivindicaciones precedentes, **caracterizado porque**, según dicha operación de selección de dicho nombre de carpeta inicial, dicho nombre de carpeta inicial es seleccionado comenzando con el nombre de carpeta correspondiente a la carpeta de la rama más exterior de dicha ruta y retrocediendo jerárquicamente en dicha ruta un número de nombres de carpeta igual a dicho número predefinido *n*.

**11.** Un producto de programa informático que comprende una porción de código de soporte lógico almacenada en la memoria de un dispositivo electrónico de procesamiento para implementar las operaciones del método según cualquiera de las reivindicaciones 1 a 10 cuando dicha porción de código de soporte lógico es ejecutada en dicho dispositivo electrónico de procesamiento.

15 **12.** Un equipo informático que comprende al menos un dispositivo electrónico de procesamiento, dotado de:

- uno o más soportes de almacenamiento, en los que están almacenados uno o más documentos digitales y en los que está cargada la porción de código de soporte lógico del producto de programa informático según la reivindicación 11;
  - un procesador configurado para ejecutar dicha porción de código de soporte lógico de tal modo que implemente las
- 20 operaciones del método según cualquiera de las reivindicaciones 1 a 10.

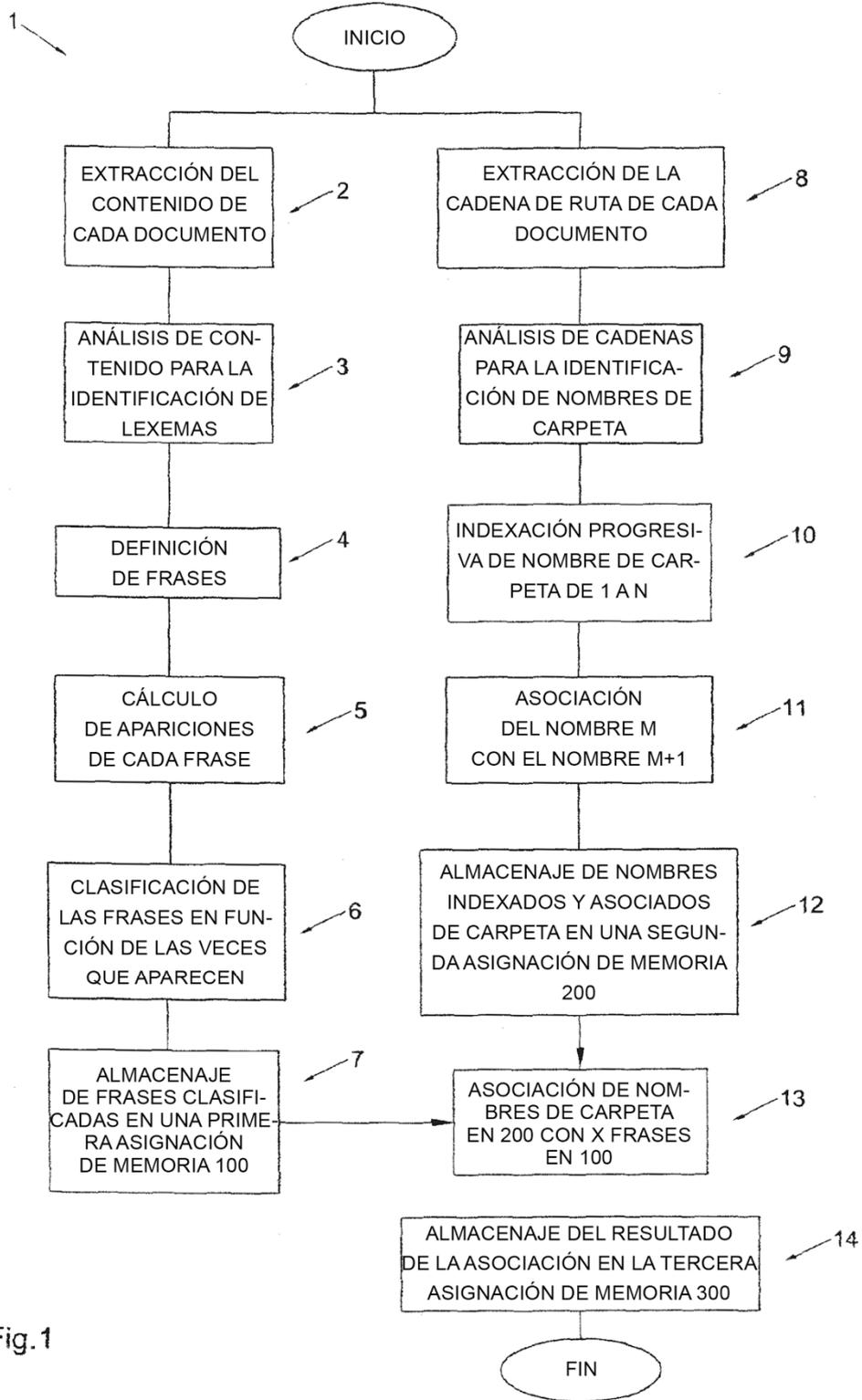


Fig.1