

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 729 488**

51 Int. Cl.:

G06F 17/30 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.05.2017 E 17173042 (7)**

97 Fecha y número de publicación de la concesión europea: **06.03.2019 EP 3249557**

54 Título: **Procedimiento implementado por ordenador y controlado por ordenador, producto de programa informático y plataforma para disponer datos para su procesamiento y almacenamiento en un motor de almacenamiento de datos**

30 Prioridad:

27.05.2016 NL 2016846

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

04.11.2019

73 Titular/es:

**DYNACTIONIZE N.V. (100.0%)
Schillerstraat 8
2050 Antwerp, BE**

72 Inventor/es:

BRANDS, MICHAEL RIK FRANS

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 729 488 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

Procedimiento implementado por ordenador y controlado por ordenador, producto de programa informático y plataforma para disponer datos para su procesamiento y almacenamiento en un motor de almacenamiento de datos

Campo técnico

- 5 La presente invención se refiere a procesamiento de datos y almacenamiento de datos, más en particular, a procesamiento implementado por ordenador y controlado por ordenador y al almacenamiento de datos legibles por ordenador de acuerdo con un formato de estructuración de datos.

Antecedentes

- 10 En su forma más general, una base de datos hace referencia a un conjunto de elementos de datos y el modelo de datos mediante el cual están interrelacionados estos elementos de datos. En bases de datos convencionales, el significado o valor de cada elemento de datos se determina por su posición en el modelo de datos. Un elemento de datos que representa el nombre dado de una persona en la base de datos obtiene su propio significado debido a que se almacena bajo, o se asigna a una etiqueta tal como 'Nombre_dado' en una tabla denominada 'Personas', por ejemplo. Es por esta razón que en los sistemas de base de datos convencionales la primera etapa en la
15 implementación y desarrollo de la base de datos es producir un modelo de datos conceptual que refleja la estructura completa de la información a mantenerse en la base de datos.

- En bases de datos convencionales, un elemento de datos puede no almacenarse si no se proporciona un encabezamiento o etiqueta específicos a los que puede asignarse el elemento de datos. Para decirlo de manera diferente, un elemento de datos puede almacenarse únicamente si un espacio particular para almacenar el elemento de datos se direcciona de antemano.
20

- En un entorno de aplicación dinámico, con conjuntos de datos grandes y grandes a almacenarse, crear tales modelos de datos 'completos' es difícil, lleva tiempo y requiere adaptaciones frecuentes, a medida que crece con el tiempo fácilmente el número de elementos que representan un significado nuevo y no previsto de un elemento de datos. Con la introducción de la telefonía móvil, por ejemplo, surgió la necesidad de expandir el modelo de datos por el
25 almacenamiento adicional de un número de teléfono móvil cerca de un número de teléfono de línea terrestre. Los expertos en la materia apreciarán las dificultades que se encuentran con la expansión de los elementos a almacenarse en un modelo de datos convencional, tal como, pero sin limitación, inconsistencias de datos, modelos ambivalentes, multiplicación de datos y así sucesivamente, todos puntos dañinos potenciales de fallo.

- Habiendo producido un modelo de datos conceptual, la siguiente etapa es traducir este modelo en una forma que realmente implemente los elementos relevantes en la base de datos. Este procedimiento a menudo se denomina el diseño de la base de datos lógica, y la salida es un modelo de datos lógico expresado en forma de un esquema. Mientras que el modelo de datos conceptual es (en teoría al menos) independiente de la elección de la tecnología de base de datos, el modelo de datos lógico se expresará en términos de una tecnología de base de datos particular.
30

- En la actualidad, el modelo de base de datos más popular para bases de datos de fin general es el modelo relacional, que usa un formato basado en tablas. El procedimiento de creación de un diseño de base de datos lógica usando este modelo implica un enfoque metodológico conocido como normalización. El objetivo de la normalización es asegurar que cada 'elemento' elemental únicamente se registra en un lugar, de modo que las inserciones, actualizaciones y borrados automáticamente mantienen la consistencia.
35

- Además del modelo relacional, y sin tener por objetivo ser completos, otros modelos de tipos de bases de datos conocidos están diseñados por un modelo de base de datos jerárquico, un modelo de red, un modelo de objetos, un modelo documental, un modelo de serie, y un modelo semántico, por ejemplo.
40

- Los datos que residen en un campo fijo dentro de un registro o fichero también se denominan datos estructurados. Es decir, los datos contenidos en bases de datos relacionales y hojas de cálculo, por ejemplo. La información que no puede clasificarse fácilmente y no se adapta a ningún recuadro particular o base de datos de fila-columna tradicional, se denomina datos no estructurados. Ejemplos de datos no estructurados son fotografías, imágenes gráficas, presentaciones, correos electrónicos, y documentos de procesamiento de textos, por ejemplo. Los ficheros no estructurados a menudo incluyen contenido de texto y multimedia. Obsérvese que aunque estos tipos de ficheros pueden tener una estructura interna, pueden considerarse aún 'no estructurados' puesto que los datos que contienen no se adaptan perfectamente en una estructura de modelo de datos rígida.
45

- Los datos semi-estructurados son una mezcla de datos estructurados y no estructurados. Son un tipo de datos estructurados, pero carecen de la estructura de modelo de datos estricta. Con datos semi-estructurados, se usan etiquetas u otros tipos de marcadores para identificar ciertos elementos dentro de los datos, pero los datos no siguen una estructura rígida. Por ejemplo, el software de procesamiento de textos puede incluir ahora metadatos que muestran el nombre del autor y la fecha de creación, mientras que el volumen del documento es solo texto no
50 estructurado.
55

Aunque un modelo de base de datos particular puede ser óptimo para almacenar uno u otro tipo de datos, en la práctica, las técnicas de modelación de base de datos conocidas todas sufren hasta un mayor o menor punto de los problemas implicados con la expansión en la cantidad y nuevos tipos de datos que se crean con el tiempo en una organización que evoluciona dinámicamente.

5 Cuando se consulta una base de datos relacional, por ejemplo, que es cuando se realiza una solicitud para recuperar información almacenada en la base de datos, no pueden revelarse relaciones entre elementos de datos distintos de los definidos por la estructura del modelo de datos. Además, muchos sistemas de base de datos requieren realizan solicitudes de información en forma de una consulta estilizada que debe escribirse en un lenguaje de consultas especial. Este es el procedimiento más complejo puesto que fuerza a los usuarios a aprender un lenguaje
10 especializado.

La solicitud de patente publicada de los Estados Unidos 2007/0106767 A1 desvela un servidor de base de datos que puede añadir diversas clases de datos asociados sin modificar la estructura de la base de datos. El servidor opera con un modelo de datos en el que los elementos de datos recibidos en un formato legible por ordenador están estructurados en una nota de encabezamiento (nodo T) y un nodo asociado (nodo A) y se define un papel de asociación (R) entre el
15 nodo T y el nodo A asociado con él. Adicionalmente, cada uno del nodo T y el nodo A se añade con un identificador (ID), su atributo (tipo de nodo), y un nombre (nombre de nodo). Los datos que tienen una estructura de este tipo se almacenan en una tabla de asociación/papel (tabla AR) compuesta de una entrada que contiene el ID del nodo A y nodo T asociados entre sí y el papel de la asociación R definido entre ellos y en una tabla de identificador (tabla de ID) que contiene el ID, atributo y nombre del nodo A y el nodo T, de modo que pueden buscarse los datos.

20 **Sumario**

Es un objeto de la presente invención proporcionar un procedimiento controlado por ordenador y una plataforma de procesamiento de datos versátiles para disponer datos legibles por ordenador estructurados, no estructurados o semi-estructurados para almacenamiento de los mismos, procedimiento y plataforma que pueden hacer frente al aumento de los datos producidos, tanto en términos de cantidad como de contenido de los mismos.

25 Es otro objeto de la presente invención proporcionar un procedimiento implementado por ordenador y controlado por ordenador y una plataforma de procesamiento de datos para la recuperación de datos almacenados de acuerdo con este procedimiento de almacenamiento de datos versátil.

En un primer aspecto se proporciona un procedimiento implementado por ordenador y controlado por ordenador para disponer datos para su procesamiento y almacenamiento de los mismos en un motor de almacenamiento de datos,
30 en el que un ordenador realiza las etapas de:

- recibir datos en un formato legible por ordenador, comprendiendo los datos una secuencia de elementos de datos;
- identificar elementos de datos en los datos recibidos y asignar a cada uno de los elementos de datos identificados un identificador de elemento de datos único que representa un respectivo elemento de datos, caracterizado porque el ordenador realiza las etapas adicionales de:
- 35 - asignar, a los datos recibidos, una acción de una pluralidad de acciones representadas por un respectivo identificador de acción único y un identificador de topología de acción único que especifica una asociación entre elementos de datos de una acción de acuerdo con una respectiva topología comprendida de una pluralidad ordenada de categorías de datos que incluyen una categoría de datos de sujeto, una categoría de datos de objeto, una categoría de datos espaciales y una categoría de datos temporales, representándose la topología por un
40 identificador de topología único y representándose cada categoría de datos por un respectivo identificador de categoría de datos;
- hacer coincidir elementos de datos identificados con la combinación de topología de acción asignada de acuerdo con el orden de la topología de acción y la secuencia de los elementos de datos, de manera que un elemento de datos se hace coincidir con una categoría de datos de la topología;
- 45 - complementar información de instancia a las combinaciones de topología de acción coincidentes y asignar un identificador de instancia de topología de acción a la combinación de topología de acción complementada, y
- almacenar, en un formato legible por ordenador, en el motor de almacenamiento de datos los elementos de datos identificados, la información de instancia y asociaciones entre identificadores resultantes de las etapas de identificación, asignación, coincidencia y complementación.

50 En lugar de tener que predefinir un esquema completo, que consiste en tablas con filas y columnas y enlaces o documentos con campos y valores, la presente invención opera de acuerdo con un formato de estructuración de datos esencialmente comprendidos de acciones, topologías, combinaciones de topología de acción, categorías de datos e información de instancia.

La invención permite que los usuarios procesen datos de la misma manera que se usan al igual que en su día cotidiano, operando piezas de datos únicas, denominadas elementos de datos, que son de interés para el usuario. La invención está basada en la idea de que tales elementos de datos - elementos de datos ya sean estructurados, no estructurados o semi-estructurados - cada uno puede estar categorizado en una o más categorías de datos de un número limitado de categorías de datos, comprendiendo al menos una categoría de datos de sujeto, una categoría de datos de objeto,
55

una categoría de datos espaciales y una categoría de datos temporales.

Es decir, la categoría de datos de sujeto hace referencia a elementos de datos recuperados por una consulta de tipo 'quién'. Elementos de datos que típicamente hacen referencia a una propiedad, es decir que hacen referencia a una capacidad o calidad y similares, han de asignarse a la categoría de datos objeto. Elementos de datos de la categoría de datos de objeto se recuperan típicamente por una consulta de tipo 'cuál'. La categoría de datos espaciales comprende elementos de datos que hacen referencia a una posición geográfica, un lugar, un espacio o similares y se recuperan por una consulta de tipo 'dónde'. Los elementos de datos que hacen referencia en general a un tiempo han de categorizarse en la categoría de datos temporales, y son el resultado de una consulta de tipo 'cuándo' aplicada en los datos a almacenarse.

5 Los elementos de datos están enlazados en el nivel de acción. Una acción específica una asociación entre elementos de datos y puede definirse como una actividad intencionada, intencional, consciente y subjetivamente significativa. En general una acción implica una intención y un objetivo. Registrar clientes de una compañía, registrar quejas, registrar los constituyentes de un producto, organizar/atender a una reunión, rellenar una tabla, actuar en una profesión, son solamente unos pocos ejemplos de acciones.

15 Aunque los elementos de datos pueden existir de manera aislada, en un escenario de uso, sin embargo, los elementos de datos se combinan siempre o están asociados con otros elementos de datos. Es una asociación de elementos de datos de este tipo en una acción la que transporta la semántica completa de cada elemento de datos para un uso particular, y tal asociación se expresa por una topología. Una topología está comprendida de una pluralidad ordenada de categorías de datos. Eso es una secuencia y números específicos de quién, cuál, dónde y cuándo. El número de categorías de datos implicado se define por la longitud de una topología. Cada topología es única en el orden de aparición y el número de las categorías de datos de una respectiva topología.

20 Por ejemplo, 'lanzar' es una acción que puede implicar que una persona, es decir un sujeto, que realiza el lanzamiento, es decir un 'quién', un objeto que ha de ser lanzado, tal como una pelota por ejemplo, es decir un 'cuál', un lugar donde se lanza la pelota, tal como en el campo de deportes, es decir un 'dónde', posiblemente un receptor de la pelota, es decir de nuevo un sujeto expresado por la categoría de datos 'quién', y tal vez el día que se lanzó la pelota, es decir un 'cuándo'. La plantilla de topología para esta acción específica está constituida entonces por las categorías de datos | quién, cuál, dónde, quién, cuándo |. Se apreciará que cuando el receptor de la pelota no es de interés, la acción implica una topología diferente, en concreto | quién, cuál, dónde, cuándo |.

25 Una topología como tal no es específica para una acción. Acciones diferentes pueden compartir una misma topología. Sin embargo, algunas acciones que implican topologías diferentes se distinguen entre sí como combinaciones de topología de acción específicas. Se apreciará que el número de acciones es virtualmente ilimitado y depende del número de sustituyentes de categoría y el número y longitud de las plantillas de topología.

30 Las combinaciones de topología de acción pueden necesitar descripción adicional que puede no estar cualificada como parte de una acción que pueda distinguirse por una topología específica, tal como que la categoría de datos de 'dónde' de una topología particular hace referencia a una ciudad o hace referencia a coordenadas geográficas, tal como datos del Sistema de Posicionamiento Global, GPS, o en que la categoría de datos 'cuándo' hace referencia a 'años' o 'segundos', por ejemplo. Sin embargo, puede complementarse otra información descriptiva a una combinación de topología de acción, como una indicación de tiempo que indica el tiempo de coincidencia, por ejemplo. Las combinaciones de topología complementadas por información descriptiva se denominan como instancias de topología de acción.

35 La invención implica el almacenamiento de los elementos de datos de los datos recibidos en asociación con al menos una categoría de datos y al menos una acción. Esto permite de manera eficaz el almacenamiento de cada elemento de datos solamente una vez, aunque un respectivo elemento de datos puede verse implicado con acciones diferentes. Proporcionando de esta manera ahorros de espacio de almacenamiento enormes, y evitando de manera eficaz duplicaciones e inconsistencias en los elementos de datos almacenados.

40 Para este fin, los elementos de datos identificados se hacen coincidir con la combinación de topología de acción asignada de acuerdo con la respectiva topología de acción y la secuencia de los elementos de datos, de manera que un elemento de datos se mapea a una categoría de datos de la topología.

45 De acuerdo con la invención, se representan acciones por un respectivo identificador de acción único. Las topologías se representan por un respectivo identificador de topología único. Una combinación de topología de acción se representa por un identificador de topología de acción único, y las instancias de topología de acción se representan por un identificador de instancia de topología de acción único. Cada categoría de datos se representa por un respectivo identificador de categoría de datos, es decir un identificador de categoría de datos de sujeto, un identificador de categoría de datos de objeto, un identificador de categoría de datos espaciales, y un identificador de categoría de datos temporales. Los elementos de datos se representan por un identificador de elemento de datos único.

50 Usando los respectivos identificadores cuando se invoca la coincidencia de los elementos de datos, la invención crea asociaciones entre los identificadores de elemento de datos, identificadores de acción, identificadores de topología de acción, identificadores de topología, identificadores de categoría de datos e identificadores de instancia de topología

de acción. Los enlaces entre elementos de datos, que en bases de datos clásicas tienen que predefinirse y expresarse por enlaces de tabla o herencia de objetos, salen a la luz en la invención por el hecho de que un mismo elemento de datos puede tomar parte en diferentes acciones. Una vez que un elemento de datos se vuelve parte de una o más acciones queda automáticamente enlazado con otros elementos de datos almacenados, independientemente del origen del elemento de datos. De hecho, no es necesario que se tenga conocimiento de antemano acerca de cualquier elemento de datos almacenado, para recuperar información de él.

Se apreciará que elementos de datos, es decir sus identificadores, pueden hacerse coincidir directamente con una categoría de datos particular, es decir un identificador de categoría de datos, de una manera predefinida, por ejemplo en caso de una única acción. De manera que la etapa de asignación se realiza de manera implícita por tal coincidencia directa.

Almacenando, en un formato legible por ordenador, en un motor de almacenamiento de datos los elementos de datos identificados en conexión con su respectivo identificador de elemento de datos, la información de instancia en conexión con un respectivo identificador de instancia de topología de acción, y las asociaciones entre los identificadores resultantes de las etapas de identificación, asignación, coincidencia y complementación, la invención permite consultar elementos de datos de la manera más flexible, simplemente a partir de la manipulación de los respectivos identificadores independientemente de los respectivos elementos de datos.

La invención separa completamente los datos sin procesar, los datos descriptivos o de instancia y los identificadores. Esto significa que la resolución de consulta puede hacerse sin datos. Incluso más, esta separación completa añade protección potente o seguridad a los elementos de datos, puesto que los elementos de datos, la información de instancia y los identificadores de enlace pueden extenderse entre hardware físicamente separado, es decir motores de almacenamiento de datos, motores de consulta de datos y otros motores de análisis de datos.

Solamente con los resultados de una consulta o un análisis de que los respectivos elementos de datos han de revelarse y a este nivel también un propietario de datos puede decidir qué elementos de datos pueden revelarse y cuáles no. La invención reduce la manipulación de grandes conjuntos de datos de manera objetiva a manipulaciones en identificadores anónimos, proporcionando de esta manera privacidad óptima.

El análisis de los identificadores como tal puede proporcionar información sobre la naturaleza, y validez de los mismos y por lo tanto de los datos almacenados, incluso sin tener que revelar elementos de datos particulares, es decir valor o significado.

En una realización la etapa de almacenamiento comprende una construcción, por el ordenador, en el motor de almacenamiento, una pluralidad de series n-dimensionales binarias, en particular una pluralidad de series bidimensionales, teniendo cada serie un único nombre e índices formados por respectivos identificadores en el que los identificadores se representan preferentemente por identificadores numéricos, y en el que las asociaciones entre identificadores se expresan por un valor binario, de manera que la presencia de una asociación entre identificadores de una serie se expresa por un primer valor binario y la ausencia de una relación entre los identificadores de una serie se expresa por un segundo valor binario.

Con un conjunto limitado de, por ejemplo, matrices de bits, pueden representarse todos los enlaces entre todos los elementos de datos posibles. Estas matrices pueden dividirse fácilmente en submatrices de tamaño fijo, que pueden almacenarse a través de un número virtualmente ilimitado de servidores que componen el motor de almacenamiento, sin la necesidad de mecanismos de reducción de mapa complejos o similares. El hecho de que este conjunto limitado - menos de aproximadamente 30 - matrices permite representar todos los enlaces en los datos también significa que los índices no crecen exponencialmente en tamaño o complejidad simplemente porque los datos se vuelvan más complejos.

La sobrecarga de cada elemento de datos único y de cada información descriptiva es mínima - fijada al tamaño de su identificador, ID, y el número de bits que hacen referencia al ID en las matrices. Las matrices de bits pueden almacenarse y recuperarse fácilmente en un disco u otro medio de almacenamiento en un formato optimizado para el espacio.

Para este fin, en una realización, los valores binarios se asignan por el ordenador para obtener series o matrices dispersas, que son matrices de las cuales el número de ceros es mayor, requiriendo menos espacio de almacenamiento. Se apreciará que esto eventualmente puede implicar la inversión de los bits que representan un enlace entre los índices de la serie o matriz.

En una realización, el ordenador almacena los identificadores de elemento de datos y los correspondientes elementos de datos, así como los identificadores de instancia de topología y la correspondiente información de instancia en listas. Estas listas, como las matrices de bits o series que expresan enlaces entre los identificadores, pueden dividirse en sublistas para almacenamiento a través de un número de servidores que componen el motor de almacenamiento. Añadiendo de esta manera otra capa de seguridad al almacenamiento de datos.

La información de instancia, que comprende información descriptiva, en particular información descriptiva que pertenece a al menos una categoría de datos de una combinación de topología, puede estructurarse de acuerdo con

el formato de estructuración de datos de la invención usado para los elementos de datos, que están comprendidos de acciones, topologías, combinaciones de topología de acción, categorías de datos e información de instancia.

5 La información de instancia, en particular, puede comprender una denominada acción de restricción y una denominada acción de aplicación. Una acción de restricción comprende un requisito que ha de satisfacerse en conexión con un elemento de datos particular. Por ejemplo, para tal elemento de datos particular una instancia o información de instancia ha de estar disponible o definida. Una acción de aplicación hace referencia a la creación de datos o activación de un evento o similares en conexión con un elemento de datos particular.

10 Estructurando la información de instancia también se asignan respectivos identificadores de la manera que se ha desvelado anteriormente en relación con los elementos de datos, se apreciará que también pueden revelarse los enlaces entre la información de instancia, proporcionando una capa adicional de análisis.

15 Un elemento de datos de acuerdo con la presente invención es una pieza de datos tal como una palabra en una frase, una parte codificada o muestra de una instantánea o un sonido, un valor de sensor, y así sucesivamente. En general, representando información o contenido de un texto, una instantánea, sonido, etc. Los elementos de datos de los datos digitales a almacenarse pueden identificarse comparando los mismos con elementos de datos almacenados en un repositorio de elementos de datos. Sin embargo, los elementos de datos pueden identificarse también por el ordenador a partir de información recibida de entrada de usuario manual o entrada remota de otras fuentes de datos. Un ejemplo sencillo son datos en los que los elementos de datos se recopilan un orden prescrito específico a través de una interfaz de usuario o menú, o separado por comas, o de otra manera. Los elementos de datos pueden recuperarse también por el ordenador aplicando unos elementos de datos que identifican un algoritmo en relación con los datos digitales a almacenarse. Unos elementos de datos que identifican el algoritmo están, por ejemplo, basados en recuento de elementos de datos estadísticos, reconocimiento de patrón, correlación y algoritmos de detección de conceptos.

20 Una acción a asignarse puede predefinirse y está disponible para el ordenador de un repositorio de acciones, por ejemplo. Sin embargo, pueden asignarse también acciones dinámicamente por entrada de usuario al ordenador, generadas por un algoritmo de identificación de acción, y/o basándose en reconocimiento de patrones por redes neurales, por ejemplo, sin sufrir los problemas de expansión del modelo de datos clásico para hacer frente a nuevos tipos y enlaces entre elementos de datos a almacenarse.

30 Una combinación de topología de acción puede asignarse, por el ordenador, recuperando una topología predefinida de topologías predefinidas almacenadas en un repositorio de topologías, por ejemplo, o por una topología recibida por el ordenador a través de entrada de usuario, por ejemplo, o de un algoritmo de identificación de topología adecuado, y combinando el mismo con la acción ya asignada. Los datos que representan un texto, una sentencia, una cláusula, una escena, un evento o similares pueden estructurarse por una o una pluralidad de topologías.

Los identificadores únicos pueden asignarse por el ordenador de acuerdo con un algoritmo adecuado como es conocido en general por un experto en la materia.

35 En el procedimiento implementado por ordenador y controlado por ordenador, se realiza una consulta de datos por un ordenador en identificadores almacenados en el motor de almacenamiento de datos.

Como se ha desvelado anteriormente, los elementos de datos están enlazados en el nivel de acción, comprendido por al menos una instancia de una respectiva topología, es decir una combinación de topología de acción, y estos enlaces se expresan por las asociaciones entre los respectivos identificadores resultantes de las etapas de identificación, asignación, coincidencia y complementación de acuerdo con el primer aspecto de la invención.

40 Por consiguiente, los elementos de datos almacenados, ya sea en solitario o en combinación, puede recuperarse, de acuerdo con la invención, realizando una consulta de recuperación de datos en el nivel de identificador.

Cualquier elemento de datos puede consultarse y todos los enlaces entre cualesquiera elementos de datos pueden descubrirse por tres tipos de consulta:

- 45 - una consulta en-acción, que identifica la coexistencia de elementos de datos en una misma acción, basándose en los identificadores de elemento de datos, identificadores de topología de acción, identificadores de acción e identificadores de topología;
- una consulta de acción común, que identifica elementos de datos que tienen un conjunto común de acciones, basándose en los identificadores de elemento de datos, identificadores de topología de acción, e identificadores de acción, y
- 50 - una consulta de correlación, que verifica elementos de datos que se comparten por topologías de acción diferentes, basándose en los identificadores de elemento de datos, identificadores de topología de acción, identificadores de acción, e identificadores de topología.

55 En una realización operando en las representaciones de serie o matriz de las asociaciones entre identificadores, se genera una matriz de consulta basándose en identificadores de elemento de datos de elementos de datos en una consulta y matrices de expresión binaria que expresan una relación lógica entre los identificadores de elemento de datos en la acción, topología y nivel de instancia de topología de acción. La consulta se resuelve solapando la matriz

de consulta con cualquiera de las series n-dimensionales o matrices que expresan una asociación entre identificadores, en el que los identificadores de elemento de datos de elementos de datos en una consulta se recuperan de elementos de datos almacenados e identificadores de elemento de datos correspondientes.

5 Aplicar operaciones de lógica de bits disponibles para un ordenador al solapar la matriz de consulta y cualquiera de las series o matrices n-dimensionales, una consulta o cuestión de sea cual sea la complejidad puede resolverse con casi velocidad idéntica.

Además de consultar o buscar elementos de datos para unos fines de análisis, la consulta de recuperación de datos puede establecerse tal que para enmascarar elementos de datos de que se recuperen para visualización selectiva y transferencia de datos, tal como se requiere en conexión con aplicaciones de seguridad de datos.

10 La estructuración de datos de acuerdo con la invención no permite únicamente una minería de datos digitales versátil y potente que incluye datos estructurados, semi-estructurados y no estructurados, para revelar la estructura e información no previstas, simplemente seleccionando o definiendo una consulta de recuperación de datos, pero también para la ejecución de operaciones.

15 Las operaciones a realizarse usando los datos almacenados, pueden ejecutarse definiendo una topología de ejecución apropiada, comprendida de categorías de ejecución. Una categoría de ejecución especifica un tipo particular de ejecución a realizarse y una topología de ejecución especifica un número y el orden en el que han de ejecutarse las respectivas categorías de ejecución, para proporcionar eventualmente una operación especificada.

20 Asignando a cada topología de ejecución un respectivo identificador de topología de ejecución único y asignando a cada categoría de ejecución un identificador de categoría de ejecución único, las operaciones pueden realizarse de una manera similar como se ha explicado anteriormente en conexión con la resolución de una consulta, solapando una matriz de ejecución comprendida de identificadores de topología de ejecución e identificadores de categoría de ejecución con cualquiera de las series o matrices n-dimensionales que expresan una asociación entre identificadores, en el que los identificadores de elemento de datos de elementos de datos implicados en una operación se recuperan de elementos de datos almacenados y correspondientes identificadores de elemento de datos.

25 El procedimiento de acuerdo con la invención es aplicable universalmente para almacenar muchos diferentes tipos de datos, tal como, pero sin limitación, datos de texto, datos lingüísticos, datos de imagen, datos de vídeo, datos de sonido, datos de control, datos de medición, datos olfativos y datos táctiles. Debido al formato de estructuración de la invención, elementos de datos de todos los tipos de datos de este tipo se almacenan y recuperan de una manera normalizada, de manera que lo mismo permite combinaciones de diferentes tipos de datos que no son posibles en absoluto con las bases de datos y estructuras de bases de datos convencionales actualmente conocidas.

30 En un segundo aspecto, se proporciona una plataforma de procesamiento de datos, configurada para disponer datos para procesamiento y almacenamiento de los mismos en un motor de almacenamiento de datos, comprendiendo la plataforma al menos un ordenador y el motor de almacenamiento de datos operativamente configurados para realizar las etapas de:

- 35 - recibir datos en un formato legible por ordenador, comprendiendo los datos una pluralidad de elementos de datos ordenados;
- identificar elementos de datos en los datos recibidos y asignar a cada uno de los elementos de datos identificados un identificador de elemento de datos único que representa un respectivo elemento de datos, caracterizado porque el al menos un ordenador y el motor de almacenamiento están operativamente configurados para realizar las etapas adicionales de:
- 40 - asignar, a los datos recibidos, una acción de una pluralidad de acciones representadas por un respectivo identificador de acción único y un identificador de topología de acción único que especifica una asociación entre elementos de datos de una acción de acuerdo con una respectiva topología comprendida de una pluralidad ordenada de categorías de datos que incluyen una categoría de datos de sujeto, una categoría de datos de objeto, una categoría de datos espaciales y una categoría de datos temporales, representándose la topología por un identificador de topología único y representándose cada categoría de datos por un respectivo identificador de categoría de datos;
- 45 - hacer coincidir los elementos de datos identificados con la combinación de topología de acción asignada de acuerdo con la topología de acción y el orden de los elementos de datos, de manera que un elemento de datos se hace coincidir a una categoría de datos de la topología;
- 50 - complementar información de instancia a las combinaciones de topología de acción coincidentes y asignar un identificador de instancia de topología de acción a la combinación de topología de acción complementada, y
- almacenar, en un formato legible por ordenador, en el motor de almacenamiento de datos los elementos de datos identificados, la información de instancia y asociaciones entre identificadores resultantes de las etapas de
- 55 identificación, asignación, coincidencia y complementación.

La plataforma de procesamiento de datos implementada por ordenador y controlada por ordenador, en una realización de la misma, puede comprender una pluralidad de dispositivos de procesamiento de datos comunicativamente interconectados que operan una pluralidad de unas capas de procesamiento de datos, tal como comprendiendo

- una primera capa de procesamiento de datos dispuesta para proporcionar al menos una interfaz de comunicación para intercambiar datos con la plataforma de procesamiento de datos;
- una segunda capa de procesamiento de datos dispuesta para implementar y controlar el formato de estructuración de datos, y
- 5 - una tercera capa de procesamiento de datos dispuesta para proporcionar acceso a una pluralidad de dispositivos de almacenamiento de datos, para fines de almacenamiento y consulta.

La plataforma de procesamiento, es decir el ordenador o los dispositivos de procesamiento de datos, en realizaciones adicionales de la misma, está configurada para realizar el procedimiento de acuerdo con el primer aspecto de la invención, anteriormente desvelado.

- 10 Los expertos en la materia apreciarán que los dispositivos de procesamiento de datos y dispositivos de almacenamiento de datos no necesitan estar situados en una única sala de datos o similares. La plataforma está estructurada para permitir procesamiento remoto y almacenamiento de datos, en particular procesamiento e interconexión basados en web.

- 15 La plataforma de procesamiento implementada por ordenador y controlada por ordenador comprende adicionalmente al menos un ordenador configurado para realizar una consulta de datos en identificadores almacenados en el motor de almacenamiento.

- 20 La invención también proporciona, en un tercer aspecto, un producto de programa informático, que comprende medio de código de programa almacenado en unos medios legibles por ordenador, estando dispuesto el medio de código para realizar el procedimiento de acuerdo con el primer aspecto de la invención, cuando el código de programa se ejecuta por un ordenador, en particular en el que los medios de código están dispuestos para integrarse en o añadirse a una aplicación informática para ejecución conjunta del código de programa y la aplicación informática por un ordenador.

- 25 Un medio legible por ordenador puede comprender cualquiera de un medio legible por ordenador transitorio o no transitorio como es conocido para los expertos en la materia. Medio legible por ordenador no transitorio para el fin de la invención incluye pero sin limitación cualquiera de medios ópticos, magnéticos, de semiconductores de estado sólido u otros, tales como los designados Discos Compactos, CD, Discos Versátiles Digitales, DVD, memoria flash, lápices de memoria, Unidades de Disco Duro, HDD, Unidades de Estado Sólido, SDD, etc.

- 30 Las anteriormente mencionadas y otras características y ventajas de la invención se entenderán mejor a partir de la siguiente descripción detallada que hace referencia a los dibujos adjuntos. En los dibujos, números de referencia similares indican partes idénticas o partes que realizan una función u operación idéntica o comparable. Los ejemplos proporcionados son para fines de ilustración únicamente no pueden considerarse como limitantes para la invención, su uso y el alcance de protección conferido por las reivindicaciones adjuntas.

Breve descripción de los dibujos

- 35 La Figura 1 muestra, de una manera muy esquemática e ilustrativa, un diagrama de tipo entidad relación que ilustra las relaciones entre las diversas entidades de los datos que estructuran el formato, de acuerdo con la invención.

La Figura 2 muestra, de una manera esquemática e ilustrativa, en una representación de diagrama de bloques, un ejemplo de un diseño lógico de una plataforma de procesamiento de datos, de acuerdo con la invención.

Las Figuras 3-8 muestran, de una manera esquemática e ilustrativa, ejemplos de elementos almacenados en los diversos almacenes de la Figura 2 con sus respectivos ID, de acuerdo con la invención.

- 40 Las Figuras 9-12 muestran algunos ejemplos de matrices bidimensionales que ilustran enlaces entre ID de diversos elementos, de acuerdo con la invención.

La Figura 13 muestra un ejemplo de una expresión de consulta basada en matriz para su uso con las matrices bidimensionales mostradas en las Figuras 9-12 y los elementos desvelados en las Figuras 3-8, de acuerdo con la invención.

45 **Descripción detallada**

En el diagrama 10 de tipo entidad relación de la Figura 1, el número 11 de referencia representa elementos de datos recibidos por un ordenador en un formato legible por ordenador. A cada uno de los elementos de datos se asigna un identificador de elemento de datos único, *ElementoDatosID*. Además de este ID, pueden identificarse adicionalmente elementos de datos por un nombre respectivo, si fuera aplicable.

- 50 Una acción o acciones a asignarse a los datos recibidos y al identificador de acción único, *AcciónID*, que representa una acción particular, se indican por el número 12 de referencia. Las acciones pueden tener también un nombre particular.

La entidad designada por el número 13 de referencia representa topologías. Las topologías están comprendidas de

una pluralidad de categorías de datos ordenadas, y cada topología se representa por un único identificador de topología, *TopologíaID*.

5 Las categorías de datos, representadas por la entidad 14, incluyen una categoría de datos de sujeto, una categoría de datos de objeto, una categoría de datos espaciales y una categoría de datos temporales. Cada categoría de datos se representa por un respectivo identificador de categoría de datos, *CategoríaDatosID*. Nombres adecuados para las categorías de datos son 'quién', 'cuál', 'dónde' y 'cuándo', que representan respectivamente la categoría de datos de sujeto, la categoría de datos de objeto, la categoría de datos espacial y la categoría de datos temporal.

10 La entidad 15 de topología de acción especifica combinaciones de acciones de la pluralidad de acciones 12 - una acción representada por un respectivo *AcciónID* único - y una topología de la pluralidad de topologías 13 - una topología representada por un respectivo *TopologíaID* único. Cada combinación de topología de acción de este tipo se representa por un identificador de combinación de topología de acción único, *AcciónTopologíaID*, que puede estar comprendido de un respectivo *AcciónID* y *TopologíaID*.

15 Para hacer coincidir los elementos 11 de datos con una combinación 15 de topología asignada, han de observarse la secuencia de los elementos de datos, es decir su *ElementoDatosID*, y el orden de las categorías 14 de datos en una topología 13 particular. Para este fin, se proporciona una entidad 16 de identificador de orden, *OrdenID*, que representa la posición de una categoría de datos particular identificada por su *CategoríaDatosID* en una topología particular, representada por su *TopologíaID*.

20 Basándose en este *OrdenID*, puede complementarse o añadirse información de instancia de una entidad 17 de información de instancia a los elementos de acción coincidentes y categorías de datos en una topología de acción, información de instancia que se representa por un identificador de información de instancia único, *InformaciónInstanciaID*.

25 Las asociaciones de los elementos de datos de esta manera coincidentes, categorías de datos y la información de instancia proporcionados se expresan por una instancia de topología de acción particular de la topología de acción, que tiene un identificador de instancia de topología de acción, *InstanciaTopologíaAcciónID*, como se representa por la entidad 18.

30 Es decir, la entidad 18 de instancia de topología de acción representa las asociaciones entre los *ElementoDatosID*, *InformaciónInstanciaID*, *AcciónTopologíaID*, *TopologíaID*, *CategoríaDatosID*, y el respectivo orden o posición de los mismos por el *OrdenID*. Las asociaciones entre los respectivos ID obtenidas de esta manera pueden almacenarse en una pluralidad de series o matrices y pueden formar una base para consulta de datos y para ejecución de datos, es decir ejecutar operaciones basándose en los elementos de datos disponibles.

35 En el ejemplo de la plataforma 20 de procesamiento de acuerdo con la invención como se muestra en la Figura 2, un ordenador que comprende una primera capa 21 de procesamiento está dispuesto para proporcionar la interfaz o interfaces de comunicación entre la plataforma 20 y el mundo externo. La primera capa 21 de procesamiento asegura el intercambio de datos entre la plataforma de procesamiento de datos y fuentes externas y consumidores de datos (no mostrados). Típicamente, la plataforma de procesamiento de datos comunica con otras aplicaciones basadas en ordenador mediante intercambios de datos automatizados o mediante intercambios accionados por el usuario.

40 Una vez que llega una solicitud de datos o una solicitud para almacenar datos con la plataforma 20 mediante una de las interfaces de comunicación de la primera capa 21 de procesamiento de la plataforma, una segunda capa 22 o de procesamiento de datos del ordenador está configurada para identificar elementos de datos y asignar una acción y topología apropiadas a los datos. Pueden identificarse elementos de datos a través de un algoritmo de identificación de elementos de datos adecuado, desde información proporcionada a través de la capa 21 de interfaz de usuario, u otras fuentes de datos e información aplicable. Adicionalmente, por la capa 22 de procesamiento de datos, las correspondientes categorías de datos coinciden con los elementos de datos basándose en su respectivo orden.

45 Para ejecutar el procedimiento de asignación, ordenación y coincidencia de los datos recibidos o solicitados en una consulta, por ejemplo, la capa 22 de procesamiento de datos interactúa con una tercera capa 23 o de acceso de datos del ordenador.

50 La capa 23 de acceso de datos comprende lógicamente al menos cuatro almacenes 26, 27, 28, 29 de datos lógicos diferentes. El primer almacén 26, que es el almacén de elemento de datos, está configurado para asignar un *ElementoDatosID* único a cada elemento de datos almacenado y para almacenar cada elemento de datos único al menos una vez (o con un grado de redundancia solicitado del usuario) en el almacén 26 de elemento de datos.

55 Los expertos en la materia apreciarán que el almacén 26 de elemento de datos puede como tal también consistir en diferentes sub-almacenes. La razón más común para tener almacenes diferentes será el tamaño y número de elementos de datos a almacenar. Sin embargo, también se podría crear un almacén específico para cada tipo diferente de elementos de datos a almacenar. Ejemplos de tipos de elemento de datos diferentes son datos de cadena, datos numéricos de coma flotante, datos de números enteros, datos de indicación de tiempo, todos los cuales pueden almacenarse en un almacén 31 en línea común o cada uno tener su propio almacén. Los datos mayores, como objetos binarios y muchos más típicamente se almacenarán en un almacenamiento 32 de base de objeto grande, LOB.

También se puede distribuir el almacén de elemento de datos lógico a través de múltiples sub-almacenes o repositorios o motores de almacenamiento físicos dispersados geográficamente, para evitar latencia y pérdidas de datos potenciales que pudieran tener lugar si los datos tuvieran que viajar a través de distancias físicas largas o en caso de que tuviera lugar un desastre en una de las localizaciones de almacenamiento físico, de manera que los datos ya no estarían más disponibles o se corrompieran, por ejemplo.

El segundo almacén 27 lógico es el almacén 27 de descripción de datos basado en acción. Este almacén de descripción de datos basado en acción contiene todos los elementos necesarios para describir adecuadamente los elementos de datos y las relaciones entre los mismos. Estos elementos son categorías 33 de datos, cada una con un *CategoríaDatosID* único, acciones 34, cada una con un *AcciónID* único, topologías 35, cada una con un *TopologíaID* único y una secuencia ordenada de los *CategoríaDatosID*, y combinaciones 36 de topología de acción, que también tienen cada una un *AcciónTopologíaID* único. De acuerdo con lo que se ha remarcado acerca del almacén 26 de elemento de datos anterior, también el almacén 27 de descripción de datos basado en acción puede consistir en sub-almacenes diferentes en función del tamaño de los datos que necesita almacenar o basándose en otros criterios relevantes al fin específico de la aplicación que usa la plataforma 20 de procesamiento de datos.

En lugar de recuperar acciones y topologías de un almacén o repositorio 34, 35, respectivamente, las mismas pueden introducirse por un usuario, por ejemplo, a través de la primera capa 21 o de interconexión de la plataforma de procesamiento de datos. Los algoritmos de identificación de acción y algoritmos de identificación de topología pueden ejecutarse por la segunda capa 22 o de procesamiento de datos de la plataforma 20.

El tercer almacén lógico, en concreto el almacén 28 de instancia de topología de acción, realmente establece el enlace entre la información en el almacén 27 de descripción de datos basado en acción y los elementos de datos en el almacén 26 de elemento de datos, usando los respectivos *ID*. El enlace se establece por medio de instancias de topología de acción. Una instancia de topología de acción persiste de la manera en la que se asigna una acción a una secuencia ordenada de elementos de datos haciendo coincidir su correspondiente información de topología expresada en la combinación de topología de acción única con la secuencia ordenada de elementos de datos identificados. Además, para cada instancia de topología de acción se hace persistir información de instancia específica. Esta información de instancia pueden ser cosas como un estado, información adicional acerca de los elementos de datos diferentes que pueden derivarse de y asociarse con su posición en la topología y la correspondiente categoría de datos y la combinación específica de la acción y la topología.

Como para todos los almacenes, el almacén de instancia de topología de acción puede consistir en múltiples sub-almacenes 37 diferentes en función del tamaño de los datos que necesitan almacenar otros criterios relevantes para el fin específico de la aplicación que usa la plataforma 20 de procesamiento de datos.

El cuarto almacén 29 lógico de la capa 23 de acceso de datos es el almacén de índice basado en modelo de acción. Este almacén contiene todos los índices necesarios para describir y hacer persistir todos los enlaces entre todos los elementos presentes en los otros almacenes de la capa 23 de acceso de datos, haciendo uso de sus respectivos identificadores.

Una manera particularmente ventajosa de representar y hacer persistir estos enlaces en el contexto de la invención desvelada es usando matrices de bits dispersas. Esto, debido a que cada clase particular de elemento en la plataforma 20 de procesamiento de datos puede representarse por un conjunto específico de identificadores numéricos y debido a que el número de conjuntos es muy limitado. Una matriz de bits dispersa puede definirse y rellenarse para todos los pares de conjuntos de identificador específicos, expresando para cada identificador en el primer conjunto si se enlaza o no a cada identificador en el segundo conjunto, usando valores booleanos 0 o 1, sin generar sobrecarga significativa.

Hay conjuntos específicos de identificadores para elementos de datos, acciones, categorías de datos, topologías, combinaciones de topología de acción, instancias de topología de acción y toda la información posicional necesaria que puede expresarse usando la longitud de la topología definida más larga como un límite superior.

El almacén 29 de índice basado en modelo de acción, que también puede comprender sub-almacenes 38 diferentes, crea la ventaja única de hacer posible expresar enlaces entre elementos de datos en el momento en el que se introducen a la plataforma de procesamiento de datos en lugar de tener que predefinirlos como se requiere en la mayoría de los sistemas de gestión de datos del estado de la técnica.

Otra ventaja única del almacén 29 de índice es que permite consultar todos los datos presentes en la plataforma de procesamiento de datos sin tener que conocer dónde en la plataforma los mismos se almacenan sin tener que unir explícitamente diferentes partes del modelo de descripción de datos como es común en la práctica en las plataformas de procesamiento de datos del estado de la técnica.

Además la combinación de identificadores numéricos, matrices de bits dispersas y la separación entre elementos de datos, elementos descriptivos e índices posibilita que la plataforma 20 de procesamiento de datos ejecute un número de operaciones tales como el control de calidad de datos, duplicación de la detección de datos, cálculos de similitud de datos, etcétera, en los datos en el sistema sin desvelar o tocar los mismos datos, y simplemente usando el conocimiento encapsulado en los conjuntos de identificadores y los enlaces entre los miembros de estos conjuntos expresados en las matrices de bits dispersas.

5 La cuarta 24 capa en la figura es la capa de ejecución de datos del ordenador. Esta capa posibilita que la plataforma de procesamiento de datos ejecute automáticamente operaciones que van a realizarse usando los datos almacenados y puede ejecutar una topología apropiada comprendida de categorías de ejecución tal como activar una señal, cargar una aplicación, detener la aplicación, enviar datos a un sistema externo y cualquiera semejante como se representa por el procedimiento 25 externo.

10 La Figura 3 muestra un conjunto de muestra de elementos de datos, representados por su respectivo *ElementoDatosID* como se almacena en el almacén 26 de elemento de datos de la plataforma 20 de la Figura 2. La Figura 4 muestra una muestra de acciones, representadas por sus *AcciónID* como se almacenan en o recuperan del almacén 34 de acción de la Figura 2. La Figura 5 describe la categoría de datos de sujeto, categoría de datos de objeto, categoría de datos espaciales y categoría de datos temporales, es decir representadas por los nombres Quién, Cuál, Dónde y Cuándo, respectivamente, e identificados por su *CategoríaDatosID* único, almacenado en el repositorio o el almacén 33 de categoría de datos de la Figura 2.

15 La Figura 6 es una lista de ejemplos de topologías y su *TopologíaID* único. Como se desvela, una topología es una lista ordenada de *CategoríaDatosID*. En la figura, por motivos de claridad, los *CategoríaDatosID* se representan también por sus respectivos nombres. Véase también la Figura 5.

La lista en la Figura 6 muestra ejemplos de combinaciones de acciones de la pluralidad de acciones 12 - una acción representada por un respectivo *AcciónID* único - y una topología de la pluralidad de topologías 13 - una topología representada por un respectivo *TopologíaID* único. Cada combinación de topología de acción de este tipo se representa por un identificador de combinación de topología de acción única, *AcciónTopologíaID*.

20 Un ejemplo de información de instancia de topología de acción, para un *InstanciaTopologíaAcciónID* particular, se muestra en la Figura 8. La información de instancia en este ejemplo comprende una indicación de tiempo, es decir *IndicaciónTiempo*, un estado, es decir *EstadoID*, que aplica acciones, es decir *AplicarAcciónID* y que restringe acciones, es decir *RestringirAcciónID*, y se almacena en el almacén 28 de instancias de topología de acción en la Figura 2. En la lista, un valor '0' en los respectivos *AcciónID* representan la ausencia de una respectiva acción para la posición correspondiente en la topología aplicable.

25 Los ejemplos anteriores son, por supuesto, no exhaustivos y se proporcionan simplemente para fines de ilustración, mientras que los ID se muestran como identificadores numéricos. Se apreciará que pueden usarse otros tipos de identificadores para el mismo fin.

30 Los enlaces entre los respectivos ID, expresados en matrices de bits de acuerdo con la invención, se muestran en los ejemplos de las Figuras 9-12. Estas matrices mostradas son simplemente unas pocas de una pluralidad de matrices o series multi-dimensionales que pueden establecerse basándose en la estructuración, procesamiento y almacenamiento de datos de acuerdo con la invención. En la nomenclatura de las matrices de bits, el primer elemento mencionado representa los ID de las filas y el último elemento mencionado representa los ID de las columnas. Las matrices se almacenan en el almacén 29 de modelo de índice basado de acción, como se muestra en la Figura 2.

35 La matriz de *ElementoDatosID* a *AcciónID* mostrada en la Figura 9, expresa asociaciones de elementos de datos, indicadas por su respectivo *ElementoDatosID*, enumeradas e impresas en negrita en la matriz de la columna más a la izquierda, y acciones, indicadas por su respectivo *AcciónID*, enumeradas e impresas en negrita en la fila superior de la matriz. En la matriz, un valor '1' indica que un respectivo elemento de datos está implicado en una respectiva acción, y un valor '0' indica que un respectivo elemento de datos no está implicado en una respectiva acción.

40 La matriz de *AcciónID* a *TopologíaID* mostrada en la Figura 10, expresa combinaciones de acciones, indicadas por su respectivo *AcciónID*, enumeradas e impresas en negrita en la columna más a la izquierda de la matriz, y topologías, indicadas por su respectivo *TopologíaID*, enumeradas e impresas en negrita en la fila superior de la matriz. En la matriz, un valor '1' indica la presencia de una respectiva acción y combinación de topología, y un valor '0' indica ausencia de tal combinación. Como puede observarse a partir de la matriz, una topología como tal no es específica para una acción. Por ejemplo, las diferentes acciones 1 y 2 comparten la misma topología 6. La acción 4 implica diferentes topologías 1, 2, 3, 4. Tales combinaciones de topología de acción se distinguen entre sí por las combinaciones de topología de acción específica.

45 La matriz *ElementoDatosID* a *CategoríaDatosID* mostrada en la Figura 11 indica a qué categoría de datos o categorías de datos, indicadas por su respectivo *ElementoDatosID*, enumeradas e impresas en negrita en la fila superior de la matriz, pertenece un elemento de datos particular, representado por su *ElementoDatosID*, enumerado e impreso en negrita en la columna más izquierda de la matriz. Por lo tanto, por ejemplo, el elemento 18 de datos es de la categoría 'Cuál', mientras que el elemento 22 de datos puede ser tanto un Quién como un Cuál. Véase también la Figura 5.

50 La matriz *TopologíaID* a *InstanciaTopologíaAcciónID* en la Figura 12, muestra que múltiples instancias de acción pueden asociarse con una respectiva topología. En la matriz de la Figura 12 las topologías se indican por su respectivo *TopologíaID*, enumeradas e impresas en negrita en la columna más a la izquierda de la matriz, y las instancias de topología de acción se indican por su respectivo *InstanciaTopologíaAcciónID*, enumerado e impreso en negrita en la fila superior de la matriz. De nuevo, en la matriz, un valor '1' indica una asociación y un valor '0' indica la ausencia de una asociación.

5 La representación matricial muestra la flexibilidad de la invención desvelada, como los enlaces en el contexto de la invención desvelada dan como resultado matrices de bits dispersas que requieren menos espacio de almacenamiento. Una matriz de bits dispersa o serie multi-dimensional puede definirse y rellenarse para todos los pares de conjuntos de identificador específicos, expresando para cada identificador en el primer conjunto si está enlazado o no a cada identificador en el segundo conjunto, usando valores booleanos 0 o 1, sin generar sobrecarga significativa.

Puede expresarse una consulta, como se muestra en la Figura 13, enumerando los elementos de datos que uno está buscando. Los enlaces lógicos que el usuario desea especificar entre los elementos de datos en la consulta pueden expresarse rellenando una matriz binaria sencilla para cada uno de los niveles en el modelo descriptivo (acciones, topologías, acción-topología-instancias).

10 Estas matrices de expresión binaria tienen tantas filas como elementos de datos hay en la consulta. El número de filas en las matrices de expresión binaria está fijado. La primera fila expresa la relación OR lógica, la segunda fila expresa la relación AND lógica y la tercera fila expresa el NOT lógico. Un uno binario, es decir un valor '1', en la posición correspondiente en la fila descrita significa que la consulta-resultado necesita rellenar la relación lógica entre el elemento de datos en la consulta y los otros elementos para el correspondiente nivel de consulta. Se proporciona una
15 fila adicional en la que puede especificarse el *CategoríaDatosID* para cada elemento de datos en la consulta que es de interés para el usuario. Los valores cero en esta fila significan que cualquier categoría de datos asociada con el elemento de datos especificado se considerará un resultado válido.

20 Se apreciará que en lugar de mencionar los mismos elementos de datos, tal como "Pete", "Lilly", "Antwerp", por ejemplo, la consulta-descripción también puede iniciar desde los respectivos *ElementoDatosID* de los elementos de datos, en aquellos casos donde son conocidos por adelantado o los elementos de datos mismos pueden no desvelarse por razones de seguridad o de otra manera.

25 Las consultas se resuelven fácilmente. En una primera etapa se realiza una búsqueda fácil en el almacén de elemento de datos para hallar los ID de los elementos de datos mencionados como consulta-datos. La segunda etapa es crear una matriz de consulta usando los ID resultantes de la primera etapa y el modelo lógico descrito en las matrices de expresión binaria, para generar la matriz de consulta que puede superponerse con el conjunto de matrices de índice que describen los enlaces entre los datos y todos los elementos descriptivos. En una tercera etapa el conjunto de ID resultantes de las operaciones de expansión y superposición se usa a continuación para capturar los elementos de datos reales en el resultado de la consulta.

30 La invención puede ponerse en práctica de otra manera que lo específicamente descrito en el presente documento, y las realizaciones anteriormente mencionadas y ejemplos se pretenden simplemente como una ilustración al lector experimentado.

REIVINDICACIONES

1. Un procedimiento implementado por ordenador y controlado por ordenador de disposición de datos para procesamiento y almacenamiento de los mismos en un motor (26, 27, 28, 29) de almacenamiento de datos, en el que un ordenador (21, 22, 23) realiza las etapas de:

- 5 - recibir (21) datos en un formato legible por ordenador, comprendiendo dichos datos una pluralidad de elementos de datos ordenados;
 - identificar (11; 22) elementos de datos en dichos datos recibidos y asignar a cada uno de dichos elementos de datos identificados un identificador de elemento de datos único (*ElementoDatosID*) que representa un respectivo elemento de datos,

10 **caracterizado porque** dicho ordenador (21, 22, 23) realiza las etapas adicionales de:

- 15 - asignar (15; 22), a dichos datos recibidos, una acción (12) de una pluralidad de acciones representadas por un respectivo identificador de acción único (*AcciónID*) y un identificador de topología de acción único (*TopologíaAcciónID*) que especifica una asociación entre elementos de datos de una acción de acuerdo con una respectiva topología (13) comprendida de una pluralidad ordenada de categorías (14) de datos que incluyen una categoría de datos de sujeto, una categoría de datos de objeto, una categoría de datos espaciales y una categoría de datos temporales, representándose dicha topología (13) por un identificador de topología único (*TopologíaID*) y representándose cada categoría de datos por un respectivo identificador de categoría de datos (*CategoríaDatosID*);
 20 - hacer coincidir (16; 22) dichos elementos de datos identificados con dicha combinación de topología de acción asignada de acuerdo con dicha topología de acción y dicho orden de dichos elementos de datos, de manera que un elemento de datos se hace coincidir con una categoría de datos de dicha topología;
 25 - complementar (18; 22) información (17) de instancia a combinaciones de topología de acción coincidentes y asignar un identificador de instancia de topología de acción (*InstanciaTopologíaAcciónID*) a dicha combinación de topología de acción complementada, y
 - almacenar (23), en un formato legible por ordenador, en dicho motor (26, 27, 28, 29) de almacenamiento de datos dichos elementos de datos identificados, dicha información de instancia y asociaciones entre identificadores resultantes de dichas etapas de identificación, asignación, coincidencia y complementación.

2. El procedimiento de acuerdo con la reivindicación 1, en el que dicha etapa de almacenamiento (23) comprende la construcción, por dicho ordenador (22), en dicho motor (29) de almacenamiento, de una pluralidad de series n-dimensionales binarias, en particular una pluralidad de series bidimensionales, teniendo cada serie un nombre único e índices formados por respectivos identificadores, en el que dichos identificadores se representan preferentemente por identificadores numéricos, y en el que las asociaciones entre identificadores se expresan por un valor binario, de manera que la presencia de una asociación entre identificadores de una serie se expresa por un primer valor binario y la ausencia de una relación entre dichos identificadores de una serie se expresa por un segundo valor binario, y en particular en el que dichos valores binarios se asignan por dicho ordenador (22) para obtener series dispersas.

3. El procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que dicha etapa de almacenamiento (23) comprende la construcción, por dicho ordenador (22), de al menos una de una lista de:

- identificadores de elemento de datos y correspondientes elementos de datos;
- identificadores de instancia de topología y correspondiente información de instancia.

4. El procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que dicha información de instancia comprende información descriptiva, en particular información descriptiva que pertenece a al menos una categoría de datos de una combinación de topología de acción, estando estructurada preferentemente dicha información de instancia de acuerdo con un formato de estructuración de datos comprendido de acciones, topologías, combinaciones de topología de acción, categorías de datos y respectivos identificadores únicos.

5. El procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que los elementos de datos de dichos datos se identifican por dicho ordenador (21, 22, 23) usando al menos uno de:

- elementos de datos almacenados en un repositorio de elementos de datos;
- información relacionada con la recepción de dichos datos, y
- un algoritmo de identificación de elementos de datos.

6. El procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que se asigna una acción por dicho ordenador (21, 22, 23) usando al menos uno de:

- acciones predefinidas almacenadas en un repositorio de acciones;
- una acción recibida por dicho ordenador, y
- un algoritmo de identificación de acción.

7. El procedimiento de acuerdo con la reivindicación 6, en el que se asigna una combinación de topología de acción por dicho ordenador (21, 22, 23) recuperando una topología de:

- topologías predefinidas almacenadas en un repositorio de topologías;
- topologías recibidas por dicho ordenador, y
- un algoritmo de identificación de topología.

5 8. El procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que se realiza una consulta de datos por un ordenador (21, 22, 23) sobre identificadores almacenados en dicho motor (26, 27, 28, 29) de almacenamiento de datos.

9. El procedimiento de acuerdo con la reivindicación 8, en el que dicha consulta comprende uno de:

- 10 - una consulta en-acción, que identifica la coexistencia de elementos de datos en una misma acción, basándose en dichos identificadores de elemento de datos, identificadores de topología de acción, identificadores de acción e identificadores de topología;
- una consulta de acción común, que identifica elementos de datos que tienen un conjunto común de acciones, basándose en dichos identificadores de elemento de datos, identificadores de topología de acción, e identificadores de acción, y
- 15 - una consulta de correlación, que verifica elementos de datos que se comparten por diferentes topologías de acción, basándose en dichos identificadores de elemento de datos, identificadores de topología de acción, identificadores de acción, e identificadores de topología.

20 10. El procedimiento de acuerdo con la reivindicación 8, cuando depende de la reivindicación 2 o 3, en el que se genera una matriz de consulta basándose en identificadores de elemento de datos de elementos de datos en una consulta y matrices de expresión binarias que expresan una relación lógica entre dichos identificadores de elemento de datos en nivel de acción, topología e instancia de topología de acción, y dicha consulta se resuelve solapando dicha matriz de consulta con cualquiera de dichas series n-dimensionales que expresan una asociación entre identificadores, en el que dichos identificadores de elemento de datos de elementos de datos en una consulta se recuperan de elementos de datos almacenados y correspondientes identificadores de elemento de datos.

25 11. Un procedimiento de acuerdo con cualquiera de las reivindicaciones anteriores, en el que las operaciones de ejecución se realizan por un ordenador (24) basándose en topologías de ejecución, cada una representada por un identificador de topología de ejecución único, y comprendida de categorías de ejecución, cada una representada por un identificador de categoría de ejecución único, y los identificadores almacenados en el motor (29) de almacenamiento de datos.

30 12. Una plataforma (20) de procesamiento de datos implementada por ordenador y controlada por ordenador, configurada para disponer datos para procesamiento y almacenamiento de los mismos en un motor de almacenamiento de datos, comprendiendo dicha plataforma (20) al menos un ordenador (21, 22, 23) y motor (26, 27, 28, 29) de almacenamiento de datos operativamente configurados para realizar las etapas de:

- recibir (21) datos en un formato legible por ordenador, comprendiendo dichos datos una pluralidad de elementos de datos ordenados;
- 35 - identificar (11; 22) elementos de datos en dichos datos recibidos y asignar a cada uno de dichos elementos de datos identificados un identificador de elemento de datos único (*ElementoDatosID*) que representa un respectivo elemento de datos, **caracterizado porque** dicho al menos un ordenador (21, 22, 23) y motor (26, 27, 28, 29) de almacenamiento están operativamente configurados para realizar las etapas adicionales de:
- 40 - asignar (15; 22), a dichos datos recibidos, una acción (12) de una pluralidad de acciones representadas por un respectivo identificador de acción único (*AcciónID*) y un identificador de topología de acción único (*TopologíaAcciónID*) que especifica una asociación entre elementos de datos de una acción de acuerdo con una respectiva topología (13) comprendida de una pluralidad ordenada de categorías (14) de datos que incluyen una categoría de datos de sujeto, una categoría de datos de objeto, una categoría de datos espaciales y una categoría de datos temporales, representándose dicha topología (13) por un identificador de topología único (*TopologíaID*) y representándose cada categoría de datos por un respectivo identificador de categoría de datos (*CategoríaDatosID*);
- 45 - hacer coincidir (16; 22) dichos elementos de datos identificados con dicha combinación de topología de acción asignada de acuerdo con dicha topología de acción y dicho orden de dichos elementos de datos, de manera que un elemento de datos se hace coincidir con una categoría de datos de dicha topología;
- 50 - complementar (18; 22) información (17) de instancia a combinaciones de topología de acción coincidentes y asignar un identificador de instancia de topología de acción (*InstanciaTopologíaAcciónID*) a dicha combinación de topología de acción complementada, y
- almacenar (23), en un formato legible por ordenador, en dicho motor (26, 27, 28, 29) de almacenamiento de datos dichos elementos de datos identificados, dicha información de instancia y asociaciones entre identificadores resultantes de dichas etapas de identificación, asignación, coincidencia y complementación.

55 13. La plataforma (20) de procesamiento de acuerdo con la reivindicación 12, en la que dicho ordenador (21, 22, 23) y motor (26, 27, 28, 29) de almacenamiento de datos están configurados para realizar el procedimiento de cualquiera de las reivindicaciones 2 - 8.

14. La plataforma (20) de procesamiento de acuerdo con la reivindicación 12 o 13, que comprende adicionalmente al

menos un ordenador (21, 22, 23, 24) configurado para realizar una consulta de datos en identificadores almacenados en dicho motor (26, 27, 28, 29) de almacenamiento.

- 5 15. Un producto de programa informático, que comprende medios de código de programa almacenados en un medio legible por ordenador, estando dichos medios de código de programa dispuestos para realizar el procedimiento de acuerdo con cualquiera de las reivindicaciones 1 - 11, cuando dicho código de programa se ejecuta por un ordenador, en particular en el que dichos medios de código están dispuestos para integrarse en o añadirse a una aplicación informática para ejecución conjunta de dicho código de programa y dicha aplicación de ordenador por un ordenador.

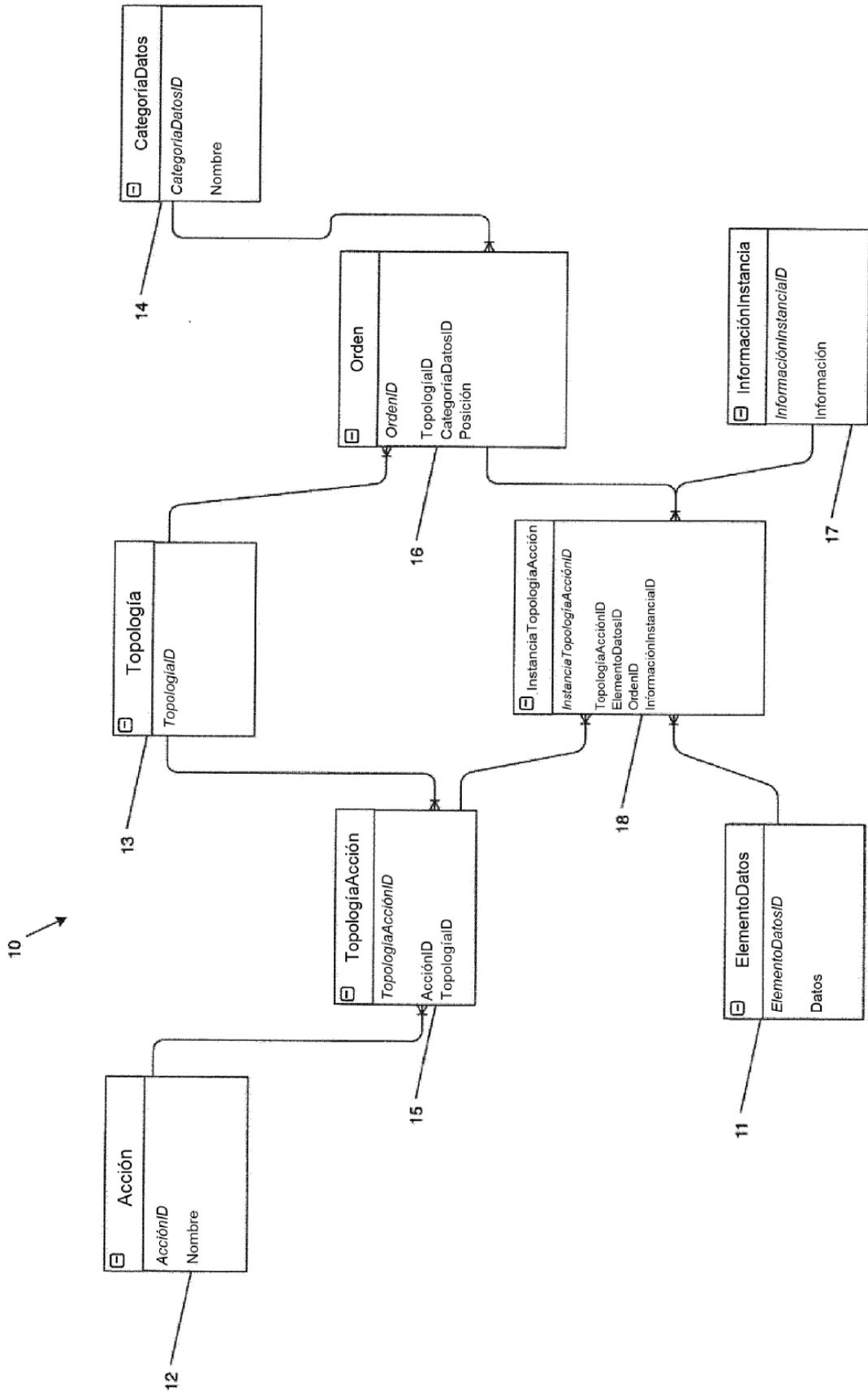


Fig. 1

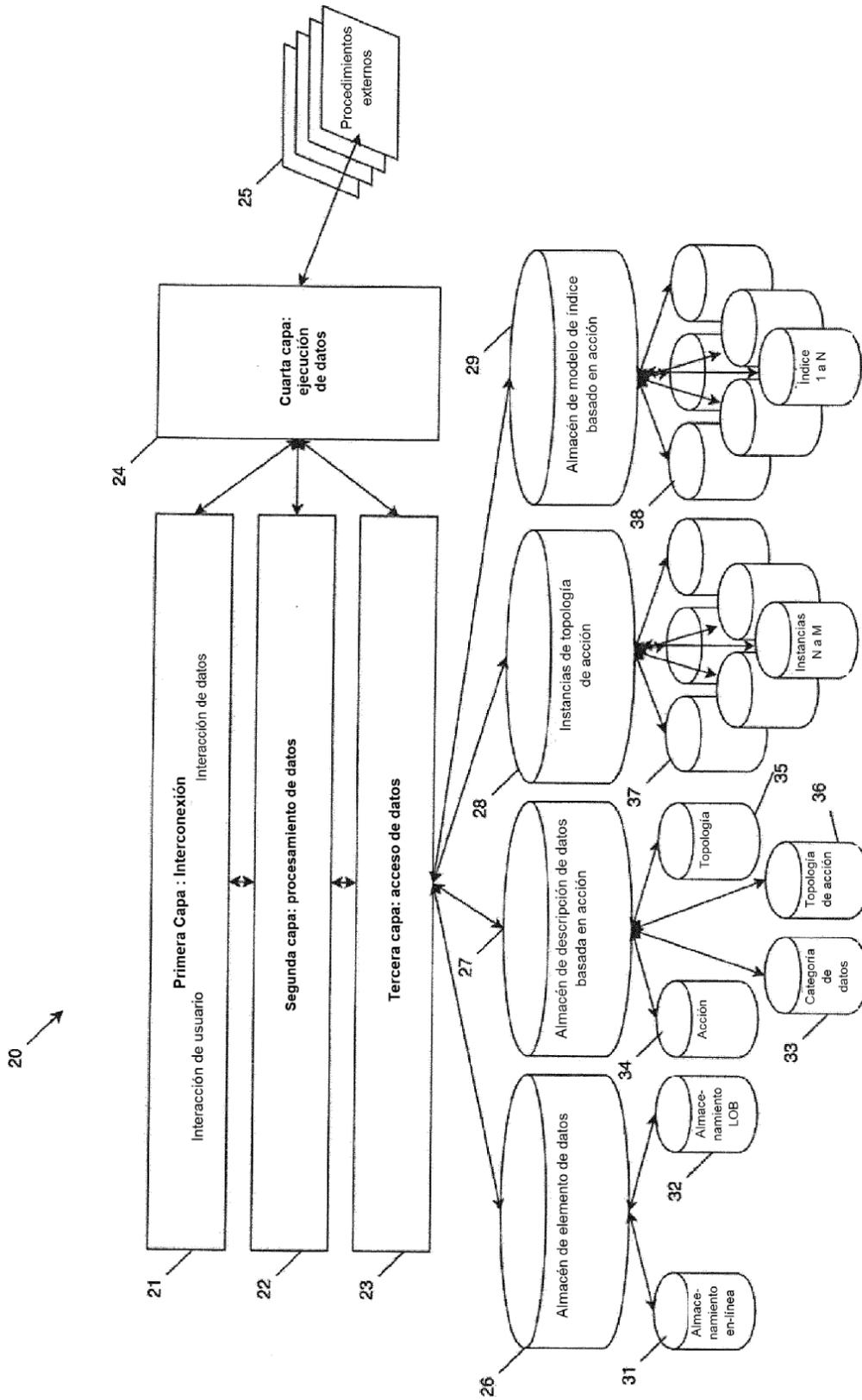


Fig. 2

ElementoDatosID	elemento de datos
...	
18	coche
19	casa
20	cafetera
21	carpintero
22	profesor
23	ciudad
24	país
25	lugar de nacimiento
...	

Fig. 3

AcciónID	NombreAcción
...	
2	HacerRecomendación
3	TieneReunión
4	DefinirConstituyente
5	PuertoDeEntrada
6	ImportaciónTabla
...	

Fig. 4

CategoríaDatosID	CategoríaDatos
1	Quién
2	Cuál
3	Dónde
4	Cuándo

Fig. 5

TopologíaID	Topología
...	
4	4,2 (CuándoCuál)
5	1,2,3 4 (QuiénCuálDóndeCuándo)
6	4,3,1 (CuándoDóndeQuién)
...	

Fig. 6

TopologíaAcciónID	AcciónID	TopologíaID
...		
5	1	6
6	5	6
7	2	7
8	2	5
...		

Fig. 7

Instancia Topología AcciónID	Indicación Tiempo	EstadoID	Topología AcciónID	Aplicar AcciónID	Restringir AcciónID	Elemento DatosID
...						
18	2016-02-02T08:18:24.928326	1	3			59,24
19	2016-02-02T08:18:24.928410	1	5	0,9,2,3		53,41,4
20	2016-02-02T08:18:24.928493	1	6		0,10,8,4	53,48,5

Fig. 8

matriz ElementoDatosID a AcciónID

	1	2	3	4	5	6	7	8	9	10
18	0	0	0	0	0	1	0	0	0	0
19	0	0	0	0	0	1	0	0	0	0
20	0	0	0	0	0	1	0	0	0	0
21	0	1	0	0	0	0	0	0	0	0
22	0	0	0	0	0	1	0	0	0	0
23	0	0	0	1	0	0	0	0	0	0
24	0	0	0	1	0	0	0	0	0	0
25	0	0	0	1	0	0	0	0	0	0

Fig. 9

matriz AcciónID a TopologíaID

	1	2	3	4	5	6	7	8
1	0	0	0	0	0	1	0	0
2	0	0	0	0	0	0	1	0
3	0	0	0	0	1	0	0	0
4	1	1	1	1	0	0	0	0
5	0	0	0	0	0	1	0	0
6	0	0	0	0	0	0	0	1

Fig. 10

matriz ElementoDatosID a CategoríaDatosID

	1	2	3	4
18	0	1	0	0
19	0	1	0	0
20	0	1	0	0
21	0	1	0	0
22	1	1	0	0
23	0	1	0	0
24	0	1	0	0
25	0	1	0	0

Fig. 11

matriz TopologíaID a InstanciaTopologíaAcciónID

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
1	0	1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	1	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	1	1	0	0	1	1	1	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0
7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
8	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Fig. 12

EXPRESIÓN DE CONSULTA

DATOS = ELEMENTODATOS ["PETE", "LILLY", "ANTWERP"]

ACCIÓN =

```

[0 1 1;      # OR
 1 0 0;      # AND
 0 0 0]      # NOT
0 0 0        # CategoríaDatosID
    
```

TOPOLOGÍA =

```

[0 1 1;      # OR
 0 0 0;      # AND
 0 0 0]      # NOT
0 0 0        # CategoríaDatosID
    
```

INSTANCIATOPOLOGÍAACCIÓN =

```

[0 1 1;      # OR
 0 0 0;      # AND
 0 0 0]      # NOT
0 0 0        # CategoríaDatosID
    
```

Fig. 13