

REIVINDICACIONES

1. Un sistema implementado por computadora para optimización de consulta e índice para recuperar datos en casos de una estructura de datos de formulación desde una base de datos, que comprende:

un dispositivo de memoria que almacena un conjunto de instrucciones; y

al menos un procesador que ejecuta el conjunto de instrucciones para llevar a cabo un método, el método comprende:

presentar una fuente de información para buscar la presencia de una o más formulaciones;

generar datos de formulación a partir de entradas de campo, en donde los datos de formulación están asociados con una o más formulaciones encontradas;

generar un caso de una estructura de datos de formulación, en donde el caso de la estructura de datos de formulación asocia la fuente de información con dichas una o más formulaciones encontradas;

crear datos de índice optimizado a partir de datos recuperados en el caso de la estructura de datos de formulación, en donde los datos de índice optimizado (i) comprenden un mapeo entre uno o más términos de campo de búsqueda potenciales y los datos de formulación, y (ii) se agrupan con base en un patrón de acceso predicho;

ejecutar una consulta de búsqueda a través de los

datos de índice optimizado; y

proporcionar información asociada con una fuente de información encontrada asociada con datos recuperados en un caso de una estructura de datos de formulación.

2. El sistema de acuerdo con la reivindicación **1**, en donde los datos de índice optimizado son un índice invertido.

3. El sistema de acuerdo con la reivindicación **1**, en donde los datos de índice optimizado se agrupan con base en un patrón de acceso predicho de tal manera que disminuya el tiempo de acceso de un motor de búsqueda de los datos de índice optimizado.

4. El sistema de acuerdo con la reivindicación **1**, en donde los datos de formulación comprenden datos de componente asociados con uno o más componentes.

5. El sistema de acuerdo con la reivindicación **4**, en donde los datos de componente comprenden datos de sustancia asociados con una o más sustancias.

6. El sistema de acuerdo con la reivindicación **5**, en donde los datos de sustancia comprenden al menos uno de un número de registro, un identificador, una tabla de conexión química, un diagrama de estructura, o un valor de propiedad numérica específica.

7. El sistema de acuerdo con la reivindicación **1**, en donde el método además comprende presentar estadísticas de búsqueda alterna.

8. El sistema de acuerdo con la reivindicación **1**, en donde el método además comprende asignar un peso de

relevancia a la fuente de información encontrada.

9. El sistema de acuerdo con la reivindicación **1**, en donde la consulta de búsqueda comprende uno o más términos de búsqueda asociados con uno o más campos de búsqueda.

10. El sistema de acuerdo con la reivindicación **9**, en donde dichos uno o más campos de búsqueda pertenecen a un campo científico.

11. El sistema de acuerdo con la reivindicación **1**, en donde dichas una o más formulaciones son formulaciones químicas.

12. El sistema de acuerdo con la reivindicación **1**, en donde los datos recuperados en un caso de la estructura de datos de formulación asociados con la fuente de información encontrada se asocian con un identificador de formulación.

13. Un medio legible por computadora no transitorio que almacena un conjunto de instrucciones que son ejecutables por al menos un procesador para llevar a cabo un método para optimización de consulta e índice para recuperar datos en casos de una estructura de datos de formulación desde una base de datos, el método comprende:

presentar una fuente de información para buscar la presencia de una o más formulaciones;

generar datos de formulación a partir de entradas de campo, en donde los datos de formulación están asociados con una o más formulaciones encontradas;

generar un caso de una estructura de datos de formulación, en donde el caso de la estructura de datos de

formulación asocia la fuente de información con dichas una o más formulaciones encontradas;

crear datos de índice optimizado a partir de datos recuperados en el caso de la estructura de datos de formulación, en donde los datos de índice optimizado (i) comprenden un mapeo entre uno o más términos de campo de búsqueda potenciales y los datos de formulación, y (ii) se agrupan con base en un patrón de acceso predicho;

ejecutar una consulta de búsqueda a través de los datos de índice optimizado; y

proporcionar información asociada con una fuente de información encontrada asociada con datos recuperados en un caso de una estructura de datos de formulación.

14. El medio legible por computadora no transitorio de acuerdo con la reivindicación **13**, en donde los datos de índice optimizado son un índice invertido y se agrupan con base en un patrón de acceso predicho de tal manera que disminuya el tiempo de acceso de un motor de búsqueda de los datos de índice optimizado.

15. El medio legible por computadora no transitorio de acuerdo con la reivindicación **13**, en donde los datos de formulación comprenden datos de componente asociados con uno o más componentes, y los datos de componente comprenden datos de sustancia asociados con una o más sustancias.

16. El medio legible por computadora no transitorio de acuerdo con la reivindicación **15**, en donde los datos de sustancia comprenden al menos uno de un número de registro,

un identificador, una tabla de conexión química, un diagrama de estructura, o un valor de propiedad numérica específica.

17. El medio legible por computadora no transitorio de acuerdo con la reivindicación **13**, en donde el método además comprende presentar estadísticas de búsqueda alterna y asignar un peso de relevancia a la fuente de información encontrada.

18. El medio legible por computadora no transitorio de acuerdo con la reivindicación **13**, en donde:

la consulta de búsqueda comprende uno o más términos de búsqueda asociados con uno o más campos de búsqueda;

dichos uno o más campos de búsqueda pertenecen a un campo científico; y

dichas una o más formulaciones son formulaciones químicas.

19. El medio legible por computadora no transitorio de acuerdo con la reivindicación **13**, en donde los datos recuperados en un caso de la estructura de datos de formulación asociados con la fuente de información encontrada se asocian con un identificador de formulación.

20. Un método para optimización de consulta e índice para recuperar datos en casos de una estructura de datos de formulación desde una base de datos, el método comprende:

presentar una fuente de información para buscar la presencia de una o más formulaciones;

generar datos de formulación a partir de entradas de campo, en donde los datos de formulación están asociados con una o más formulaciones encontradas;

generar un caso de una estructura de datos de formulación, en donde el caso de la estructura de datos de formulación asocia la fuente de información con dichas una o más formulaciones encontradas;

crear datos de índice optimizado a partir de datos recuperados en el caso de la estructura de datos de formulación, en donde los datos de índice optimizado (i) comprenden un mapeo entre uno o más términos de campo de búsqueda potenciales y los datos de formulación, y (ii) se agrupan con base en un patrón de acceso predicho;

ejecutar una consulta de búsqueda a través de los datos de índice optimizado; y

proporcionar información asociada con una fuente de información asociada con datos recuperados en un caso de una estructura de datos de formulación.