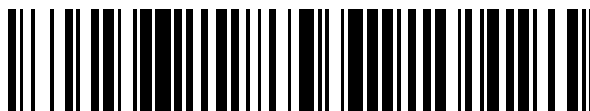


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 707 978**

51 Int. Cl.:

**G06Q 30/02** (2012.01)

**G06F 11/34** (2006.01)

**H04L 29/08** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.02.2014 PCT/FI2014/050140**

87 Fecha y número de publicación internacional: **03.09.2015 WO15128536**

96 Fecha de presentación y número de la solicitud europea: **26.02.2014 E 14884200 (8)**

97 Fecha y número de publicación de la concesión europea: **31.10.2018 EP 3111399**

54 Título: **Medición de perfiles de usuarios de Internet de múltiples pantallas, comportamientos transaccionales y estructura de población de usuarios a través de un censo híbrido y metodología de medición basada en el usuario**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**08.04.2019**

73 Titular/es:

**VERTO ANALYTICS OY (100.0%)  
Tekniikantie 14  
02150 Espoo, FI**

72 Inventor/es:

**VERKASALO, HANNU;  
MODZELEWSKI, CHRIS y  
MALMI, ERIC**

74 Agente/Representante:

**CARPINTERO LÓPEZ, Mario**

ES 2 707 978 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

Medición de perfiles de usuarios de Internet de múltiples pantallas, comportamientos transaccionales y estructura de población de usuarios a través de un censo híbrido y metodología de medición basada en el usuario

### Campo de la invención

5 En general, la invención se refiere a dispositivos digitales, comunicaciones incluyendo comunicaciones por cable y móviles, y aplicaciones relacionadas y servicios. En particular, sin embargo, no exclusivamente, la presente invención se refiere a la recopilación de hardware de Internet multipunto y datos de uso, fusionando estos datos y publicando con frecuencia métricas representativas tanto del nivel de mercado como del comportamiento de usuarios con dispositivos digitales junto con información pertinente de la estructura del mercado.

### Antecedentes

10 Varias soluciones modernas de medición y análisis de medios se centran en la recopilación de datos basados en una sola fuente. En general, las fuentes son los llamados estudios de panel en los que los comportamientos de los usuarios se miden a través de dispositivos dedicados o medidores de software descargables, o mediante etiquetas integradas (en sitios (web) o apps (aplicaciones)) o SDK (aplicaciones) que recopilan datos sobre una aplicación determinada. Alternativamente, los datos deseados se pueden adquirir a través de estudios de encuestas de usuarios tradicionales o entrevistas que sufren el problema de la subjetividad e inexactitud del entrevistado.

15 En los estudios antes mencionados, el objetivo evidente es obtener una comprensión de las tendencias subyacentes, los hábitos, los problemas y las necesidades de los usuarios. Sin embargo, cada uno de los procedimientos actuales tiene sus propios problemas subyacentes. Por ejemplo, con pocas excepciones, los costes de reclutar, mantener y validar un panel que sea representativo en términos demográficos y de comportamiento son prohibitivos. Los SDK (kit de desarrollo de software) y las etiquetas solo proporcionan datos sobre las propiedades participantes, pero no todas. Las encuestas y entrevistas son una mejor indicación de la solidez de la marca que el comportamiento real y no existe un enfoque que pueda proporcionar información conceptual sobre todas las áreas clave del ecosistema de Internet en términos de datos "duros", objetivos y basados en observaciones: base instalada en hardware y ventas, contenido y distribución de aplicaciones, y uso/transacciones completadas por el usuario.

20 La evolución de los medios de comunicación y servicios de Internet tales como sitios web o servicios web accesibles es ahora más rápida que nunca, y surgen nuevos dispositivos en el mercado continuamente. Además, un usuario normalmente tiene varios dispositivos con capacidad para Internet, en lugar de uno. Se necesitaría un entendimiento holístico no solo del uso, sino también de la distribución de dispositivos y contenido, para explicar la dinámica del mercado y para proporcionar productos de investigación todo en uno a los clientes clave que participan en el ecosistema de Internet.

25 A modo de ejemplo, la penetración de cierto tipo de dispositivos afecta a la distribución de un determinado servicio, ya sea porque está integrado de manera predeterminada en el dispositivo o tiendas de aplicaciones (u otros mecanismos de distribución de contenido) están impulsando las descargas de dicho servicio con ese dispositivo particular. Como otro ejemplo, no es suficiente comprender cuánta gente descarga ciertas aplicaciones de las tiendas de aplicaciones, pero cada vez más es importante comprender la conversión de las descargas al uso real, y además el dinero gastado por el usuario en el uso del servicio. Además, la popularidad y la variedad de servicios disponibles en un determinado dispositivo pueden afectar las ventas de ese dispositivo debido a una mejor funcionalidad percibida, la experiencia del usuario o a través de círculos sociales. Por lo tanto, hay importantes bucles de retroalimentación en este sistema.

30 Otra tendencia actual es que las personas no solo tengan múltiples dispositivos, sino que también usen múltiples interfaces de usuario, tecnologías portátiles o dispositivos conectados, todos los que trabajan total o parcialmente con el apoyo de un llamado dispositivo maestro. Ejemplos tangibles incluyen relojes inteligentes o gafas digitales del tipo de dispositivos, que se adjuntan a un dispositivo de teléfono inteligente. La medición de las actividades a través de esos dispositivos adjuntos y dispositivos portátiles ciertamente también tiene una importancia clave para el futuro.

35 Por lo tanto, existe la necesidad de una solución de medición de medios escalable capaz de realizar mediciones observacionales de Internet que se adapten mejor al entorno de los medios modernos, donde la complejidad y la fragmentación de dispositivos, aplicaciones y servicios se han convertido en la norma. Por lo tanto, una solución de medición de medios de este tipo proporcionaría un enfoque dinámico y de alta resolución para la medición y el análisis holísticos de Internet, integrando métricas en base al hardware instalado, la distribución de contenido y los comportamientos de los usuarios en un marco único, y aprovechando los mejores procedimientos de su clase para capturar cada faceta del ecosistema en línea. El resultado produce un sistema integrado y sensible capaz de proporcionar mucho más que la suma de sus partes y mucho más que las soluciones aisladas anteriores. El documento US2013/198125 divulga una solución para el tráfico web y la medición de audiencia basada en mediciones de dispositivos de panelistas y mediciones de audiencia en línea, especialmente en relación con la monitorización de contenido de medios en un sitio de visualización de contenido, en el que los datos de panelistas indicativos de acceso a páginas web en ubicaciones medidas se obtienen y se clasifican de acuerdo con la demografía de los panelistas. Se genera un panel virtual basado en las estimaciones de acceso a páginas web en

ubicaciones no medidas y se utiliza para clasificar los datos de visualización de página asociados con dichas ubicaciones no medidas.

El documento EP2341479 divulga soluciones para rastrear el uso de recursos de un usuario en Internet bajo el control del usuario, y para controlar la distribución de los datos de uso del usuario. Los puntos finales de datos personales se proporcionan en terminales de usuario conectados a una red pública. Un punto final incluye una interfaz de datos de red configurada para recibir datos comunicados hacia y desde un usuario de al menos una aplicación personal. La aplicación está configurada para enviar y recibir datos a través de la red pública. Un registrador de datos de uso almacena los datos de uso en un registro de datos de uso. Un cliente ligero operable en un dispositivo móvil puede sincronizarse con un cliente que se ejecuta en un ordenador de escritorio en términos de registro de datos de uso, utilizando, por ejemplo, la duplicación de datos.

### **Sumario**

Por lo tanto, según un aspecto de la presente invención, un sistema al menos sustancialmente electrónico para la obtención y el análisis de datos sobre el comportamiento del usuario en línea y otra actividad teniendo en cuenta los dispositivos de usuario que se pueden conectar a Internet, opcionalmente dispositivos móviles, está configurado para

recopilar datos desde una pluralidad de fuentes de datos, en el que dichos datos recopilados incluyen al menos datos individuales a nivel de usuario adquiridos desde dispositivos de usuario a través de medición pasiva y datos a nivel de censo que indican el comportamiento y las características demográficas en toda la población de usuarios, dispositivos activos o servicios medidos, siendo dichos datos a nivel de usuario preferiblemente recopilados utilizando al menos un panel de investigación de usuarios de constitución controlada, además preferiblemente desde múltiples dispositivos de cada participante en el panel de usuario, siendo dichos datos recopilados indicativos de Internet, contenido, medios, aplicación y/o uso del dispositivo,

organizar los datos a nivel de usuario obtenidos y los datos a nivel de censo en un conjunto de datos categorizados multivalentes preferiblemente utilizando un esquema de metadatos ontológicos,

determinar un entregable electrónico basado en una integración de los datos a nivel de usuario y los datos a nivel de censo, en el que los datos a nivel de censo se utilizan para calibrar los datos a nivel de usuario,

teniendo el entregable un alcance seleccionable dinámicamente, preferiblemente seleccionable por el usuario, en términos de una serie de factores de interés relacionados con dispositivos usados o comportamientos en línea, incluyendo preferiblemente el uso de aplicaciones, la distribución de aplicaciones, el uso de contenido, la distribución de contenido, la monetización de aplicaciones, los datos demográficos del usuario, la distribución de los dispositivos, las características de los dispositivos, el uso de los dispositivos y/o factores de tiempo,

en el que el sistema está configurado para recopilar dichos datos a nivel de usuario a través de uno o más dispositivos maestros de un usuario que tiene una pluralidad de dispositivos de usuario, en el que hay una serie de otros dispositivos de usuario conectados funcionalmente a dicho uno o más de dichos dispositivos maestros, midiéndose el uso de dichos otros dispositivos de usuario medidos por los dispositivos maestros correspondientes, actuando los dispositivos maestros como recopilación de datos y nodos de conexión hacia el sistema,

y, además, en el que el sistema está configurado para distinguir, en el análisis que da como resultado lo que se puede entregar, entre datos a nivel de usuario que se originan en diferentes dispositivos de usuario de la pluralidad de dispositivos de usuario del usuario en función de los identificadores de dispositivo asociados con los datos a nivel de usuario adquiridos a partir de los dispositivos del usuario.

Opcionalmente, el sistema puede incorporar además la recogida de datos de estratificación estadísticos derivados de una muestra que estima la distribución de las características a través de la población de usuarios, con lo cual el sistema está configurado para organizar también los datos de estratificación en el conjunto de datos y utilizarlos en la determinación del entregable, incluyendo la calibración de datos a nivel de usuario.

En una realización, los datos a nivel de usuario se recogen más desde un segundo panel de usuario, potencialmente mayor. Preferiblemente aún, ambos paneles comprenden naturalmente una pluralidad de usuarios. Preferiblemente, los paneles no tienen superposición (no hay usuarios compartidos). El control sobre la constitución del segundo panel de usuario puede reducirse en comparación con el primer panel. El segundo panel puede estar abierto a, si no prácticamente a todos, al menos a un gran número de participantes que pueden necesitar simplemente cumplir con una serie de requisitos básicos tal como la edad adulta legal.

En otra realización, complementaria o alternativa, la recopilación de datos que utiliza los dispositivos de usuario incluye la incrustación de software de medición relacionada en otra(s) pieza(s) de software funcionalmente independientes, incluyendo el otro software preferiblemente software que está diseñado principalmente para un propósito diferente de la medición/registro de datos y/o proporciona un beneficio distinto para el usuario en términos de, por ejemplo, funciones relacionadas con el entretenimiento, la comunicación o la mensajería, herramientas de procesamiento de datos, etc. El componente de software de medición puede incluirse como un módulo adicional en un solo dispositivo software electrónico entregable al usuario.

En una realización adicional, complementaria o alternativa, los datos a nivel de usuario se recogen a través de uno o más dispositivos maestros de un usuario, en el que los dispositivos maestros pueden ser uno o más de otros dispositivos conectados, tales como dispositivos UI (interfaz de usuario) externos (gafas inteligentes, etc.) al menos

funcionalmente conectados, cuyo uso se mide utilizando los dispositivos maestros como el punto de conexión a los servidores de recogida de datos de la red, y utilizando directamente (por ejemplo, observando el estado y la actividad del dispositivo conectado) a través de llamadas directas a la API del dispositivo conectado) o indirectamente (por ejemplo, al observar el seguimiento del tráfico entre el dispositivo conectado y la red de Internet a través de un dispositivo que funciona como un llamado dispositivo de pasarela), los dispositivos maestros registran la actividad de los dispositivos conectados.

Opcionalmente, en el caso de la medición de múltiples dispositivos, el dispositivo maestro al que está conectado al menos un dispositivo conectado, está configurado para realizar al menos una acción seleccionada del grupo que consiste en:

- a. Rastrear el tráfico, las señales u otros eventos que tienen lugar entre el(los) dispositivo(s) conectado(s) y el dispositivo maestro;
- b. Encuestar de forma pasiva o activa las interfaces de programación de la aplicación del (los) dispositivo(s) conectado(s) para recopilar datos de medición;
- c. Almacenar los datos de medición recopilados en el dispositivo maestro temporalmente o durante un período de tiempo predeterminado; y
- d. Enviar los datos a los servidores de recopilación de datos basados en la red para fines de análisis.

En una realización adicional, complementaria o alternativa, los datos a nivel censo recopilados incluyen datos, tal como un número de puntos de datos, que describen la distribución de las características entre toda la población objetivo o un subconjunto completo y rigurosamente definido de la misma, obtenida a partir de una serie de fuentes de datos centradas en la red de Internet, que pueden incluir, entre otros, datos de portadores, datos de redes publicitarias, datos de intercambio de anuncios, plataformas de análisis de sitios/aplicaciones proporcionadas, pasarela de red o datos proxy, y/o suministro de demanda u otros datos basados en plataformas de entrega de datos.

Aplicando las agregaciones de Internet (red) basadas en datos procesados y modeladas en general de una manera congruente con las descripciones proporcionadas en el presente documento, se pueden poblar las estimaciones de la base de usuarios de corriente estática de usuario de destino dispositivos (Internet), por ejemplo, en un mercado determinado. Diferentes características, tal como la tasa de crecimiento o la disminución de dicha población, pueden modelarse dinámicamente. En otras palabras, los datos a nivel de censo, tales como los datos de red (-céntricos), podrían utilizarse, potencialmente incluso como una solución independiente que omita el aspecto de calibración de datos del usuario generalmente expuesto y preferido en el presente documento, para estimar la estructura y, por ejemplo, el crecimiento de la base instalada de dispositivos de Internet.

Sin embargo, en una realización adicional, ya sea complementaria o alternativa, los datos a nivel de censo recopilados incluyen datos, tal como el número de puntos de datos, que describen los comportamientos agregados de la población objetivo, cuyos datos se pueden obtener, por ejemplo, a partir de un número de plataformas de distribución de contenido, tal como tiendas de aplicaciones públicas o lugares de mercado digital disponibles para el acceso de los usuarios a través de dispositivos digitales (usuarios) y, por ejemplo, Internet.

Opcionalmente, en el caso de la utilización de datos a nivel de censo donde los datos obtenidos de los mismos son incompletos, los datos incompletos pueden complementarse o ajustarse utilizando un número de procedimientos de modelado estadísticamente válidos, tales como la atribución, calibración, y la fusión que se ajustan a las mejores prácticas teóricas y metodológicas aplicables.

En una realización adicional, ya sea complementaria o alternativa, las fuentes de nivel de usuario y a nivel de censo de datos están integradas a través de una plataforma de superposición de metadatos, que modela interrelaciones entre conceptos, temas, y entidades observados en las fuentes de datos diferentes antes y después de la calibración y el procesado.

En una realización adicional, ya sea complementaria o alternativa, un número de dispositivos de usuario están configurados para proporcionar datos con respecto a eventos de comportamiento o contextuales externos al dispositivo, por ejemplo, para entregables de reconocimiento de sonido ambiente o de imágenes de fotos de la cámara o grabaciones de audio capturadas automáticamente, para ser enviadas junto con otros datos de medición recopilados por el software de medición a los servidores externos de recogida de datos del sistema para fines de análisis. En otro aspecto, un procedimiento para obtener y analizar datos sobre el comportamiento del usuario en línea y otra actividad relacionada con dispositivos de usuario que pueden conectarse a Internet, opcionalmente dispositivos móviles, comprende recopilar datos desde una pluralidad de fuentes de datos, en el que dichos datos recopilados incluyen al menos datos individuales a nivel de usuario adquiridos desde dispositivos de usuario a través de medición pasiva y datos a nivel de censo que indican el comportamiento y las características demográficas en toda la población de usuarios, dispositivos activos o servicios medidos, siendo dichos datos a nivel de usuario preferiblemente recopilados utilizando al menos un panel de usuarios de constitución controlada, además preferiblemente desde múltiples dispositivos de cada participante en el panel de usuario, siendo dichos datos recopilados indicativos de Internet, contenido, medios, aplicación y/o uso del dispositivo,

organizar los datos a nivel de usuario obtenidos y los datos a nivel de censo en un conjunto de datos categorizados multivalentes utilizando un esquema de metadatos ontológicos, y determinar un entregable electrónico basado en una integración de los datos a nivel de usuario y los datos a nivel de censo, en el que los datos a nivel de censo se utilizan para calibrar los datos a nivel de usuario,

5 en el que el entregable tiene un alcance seleccionable dinámicamente, preferiblemente seleccionable por el usuario, en términos de una serie de factores de interés relacionados con dispositivos usados o comportamientos en línea, incluyendo preferiblemente el uso de aplicaciones, la distribución de aplicaciones, el uso de contenido, la distribución de contenido, la monetización de aplicaciones, los datos demográficos del usuario, la distribución de los dispositivos, las características de los dispositivos, el uso de los dispositivos y/o factores de tiempo,

10 y en el que dichos datos a nivel de usuario se recopilan a través de uno o más dispositivos maestros de un usuario que tiene una pluralidad de dispositivos de usuario, en donde hay una serie de otros dispositivos de usuario conectados funcionalmente a dicho uno o más de dichos dispositivos maestros, el uso de dichos otros dispositivos de usuario medidos por los dispositivos maestros correspondientes, los dispositivos maestros que actúan como recopilación de datos y nodos de conexión hacia el sistema,

15 y, además, en el que se distingue, mediante el sistema en el análisis que da como resultado lo que se puede entregar, entre datos a nivel de usuario que se originan en diferentes dispositivos de usuario de la pluralidad de dispositivos de usuario del usuario en función de los identificadores de dispositivo asociados con los datos a nivel de usuario adquiridos a partir de los dispositivos del usuario.

20 En un aspecto adicional, un producto de programa informático realizado en un medio legible por ordenador no transitorio, comprende un código para la obtención y el análisis de datos sobre el comportamiento del usuario en línea y otra actividad, teniendo en cuenta los dispositivos de usuario que se pueden conectar a Internet, opcionalmente dispositivos móviles, incorpora

25 recopilar datos desde una pluralidad de fuentes de datos, en el que dichos datos recopilados incluyen al menos datos individuales a nivel de usuario adquiridos desde dispositivos de usuario y datos a nivel de censo que indican el comportamiento y las características demográficas en toda la población de usuarios, dispositivos activos o servicios medidos, siendo dichos datos a nivel de usuario preferiblemente recopilados utilizando al menos un panel de usuarios de constitución controlada, además preferiblemente desde múltiples dispositivos de cada participante en el panel de usuario, siendo dichos datos recopilados indicativos de Internet, contenido, medios, aplicación y/o uso del dispositivo,

30 organizar los datos a nivel de usuario obtenidos y los datos a nivel de censo en un conjunto de datos categorizados preferiblemente multivalentes utilizando un esquema de metadatos ontológicos, y determinar un entregable electrónico basado en una integración de los datos a nivel de usuario y los datos a nivel de censo, en el que los datos a nivel de censo se utilizan para calibrar los datos a nivel de usuario,

35 en el que el entregable tiene un alcance seleccionable dinámicamente, preferiblemente seleccionable por el usuario, en términos de una serie de factores de interés relacionados con dispositivos usados o comportamientos en línea, incluyendo preferiblemente el uso de aplicaciones, la distribución de aplicaciones, el uso de contenido, la distribución de contenido, la monetización de aplicaciones, los datos demográficos del usuario, la distribución de los dispositivos, las características de los dispositivos, el uso de los dispositivos y/o factores de tiempo.

40 La utilidad de la presente invención reside en una variedad de problemas, dependiendo de la realización particular en cuestión. La invención proporciona una solución factible para combinar ciertos tipos de datos centrados en el usuario y en la red, con lo cual, a través de una metodología integrada, se pueden obtener convenientemente informes electrónicos deseados y otros entregables de información, en los que la representatividad y el alcance de la información disponible es superior a la que se puede lograr utilizando diferentes fuentes de datos por separado, y la disponibilidad de métricas en el hardware y el comportamiento de los usuarios es más amplia de lo que se había visto anteriormente. La utilidad de la invención también radica en las formas en que los actores de la industria pueden interrogar dichos datos para informar de sus decisiones tácticas y estratégicas, incluidas las decisiones sobre marketing de productos, diseño de campañas de marketing, características o precios de productos, o decisiones de inversión relacionadas con jugadores individuales en el ecosistema digital.

45 Además, diferentes realizaciones de la presente invención permiten la construcción de entregables para identificar las tendencias del mercado, obtener puntos de vista de la competencia, y para conocer nuevas oportunidades emergentes en el mercado digital. La invención descrita proporciona un marco de múltiples pantallas para comprender oportunamente las diversas características del mercado digital, determinando efectivamente un sistema que puede obtener datos a nivel de censo en el mundo actual de múltiples pantallas, incluyendo los dispositivos móviles, y conectando esto a procedimientos novedosos de recopilar datos transaccionales, de comportamiento y contextuales de dichos dispositivos, procesar los datos resultantes a través de un marco de metadatos integrado y proporcionar métricas de mercado agregadas a través de un cálculo de proyección de medición de audiencia unificada.

50 Sin embargo, diversas realizaciones de la presente invención producen soluciones económica y técnicamente más escalables que la técnica anterior, y se cree que proporcionan estimaciones más precisas de una gama más amplia de métricas sobre una sección transversal más ancha del mercado digital. Debido a su diseño y dependencia en fuentes de datos multivalentes que utilizan diversos datos a nivel de usuario, diversos datos a nivel de censo y diversos datos de estratificación derivados de muestras, las realizaciones generalmente son más capaces de adaptarse a los cambios en el mercado digital, en particular, los observables en tendencias de mercado

prevalecientes en torno al desarrollo, distribución y uso de contenido digital y fabricación, distribución y uso de dispositivos.

La expresión "un número de" se refiere aquí a cualquier entero positivo a partir de uno (1), por ejemplo, a uno, dos, o tres.

- 5 La expresión "una pluralidad de" se refiere aquí a cualquier entero positivo a partir de dos (2), por ejemplo, a dos, tres, o cuatro.

La expresión "transferencia de datos" puede referirse a la transmisión de datos, la recepción de datos, o ambos, dependiendo de la(s) función(es) de una entidad particular bajo análisis relativo de una acción de transferencia de datos, es decir, un papel de un remitente, un papel de un destinatario, o ambos.

- 10 Los términos "un" y "una" no indican una limitación de la cantidad, sino que indican la presencia de al menos uno del elemento referenciado.

Los términos "primero" y "segundo" no indican ningún orden, cantidad o importancia, sino que se utilizan para distinguir un elemento de otro.

- 15 El término "censo" se refiere aquí a un sistema o conjunto de datos que refleja el comportamiento, características, o el rendimiento de preferencia de todos los miembros de un grupo dado y definido con precisión de los sujetos medidos (por ejemplo, usuarios, servicios, dispositivos, etc.). Estos datos se basan explícitamente en todos los sujetos/miembros de la población definida y medida, aunque la población definida y medida en sí misma puede ser naturalmente un subconjunto de una población más grande.

- 20 El término "calibración" se refiere aquí a un proceso de modelado estadístico mediante el cual se ponderan datos para ajustar, por ejemplo, la falta de respuesta de la muestra y otros sesgos para reflejar mejor la población medida de los usuarios, servicios o dispositivos. Este proceso de modelado estadístico puede involucrar una variedad de técnicas estadísticas bien establecidas en la literatura, que incluyen, entre otras, el balanceo de muestras, la atribución de características, la proyección de características y la integración de datos.

- 25 El término "base instalada" de dispositivos se refiere a la información y a las métricas de comunicación y que refleja el número, la estructura y/o la propiedad cruzada de dispositivos de Internet (incluyendo preferiblemente dispositivos maestros y conectados) distribuidos o en uso, por ejemplo, en un mercado dado, entre una determinada población objetivo.

- 30 El término "conjunto de datos" se refiere en este documento a una colección de elementos de datos asociados entre sí por medio de un esquema ontológico, por ejemplo. Los elementos de datos pueden ubicarse físicamente como distribuidos entre varios dispositivos y/o estructuras de datos, tal como bases de datos o tablas de datos.

Diferentes realizaciones de la presente invención se divulgan en las reivindicaciones dependientes. Varias otras realizaciones y beneficios adicionales de la invención se hacen evidentes para un experto en base a la siguiente descripción detallada.

### **Breve descripción de los dibujos adjuntos**

- 35 Se describirán varios aspectos y realizaciones de la presente invención en mayor detalle con referencia a los dibujos adjuntos, en los que

La figura 1 ilustra una realización del sistema de acuerdo con la presente invención desde un punto de vista de hardware.

- 40 La figura 2 ilustra la realización del sistema con un enfoque más conceptual y funcional en contraste con la figura 1.

La figura 3 ilustra una realización del uso de múltiples dispositivos que incorporan una serie de dispositivos maestros y dispositivos de usuario conectados.

La figura 4 proporciona una visión más detallada de las partes internas de una realización de un sistema de acuerdo con la presente invención con vistas funcionales/lógicas y de hardware.

- 45 La figura 5 ilustra diferentes poblaciones, potencial y probablemente, superpuestas, que pueden medirse mediante diferentes realizaciones de la presente invención.

La figura 6 ilustra una realización de un sitio web, o particularmente una tienda de aplicaciones, técnica de rastreo que puede utilizarse en relación con la presente invención.

- 50 La figura 7 muestra diferentes aspectos de una realización con respecto a la estimación de estadísticas de descarga, por ejemplo, número de descargas de una aplicación, en función de la información disponible en forma de, por ejemplo, clasificación de aplicaciones.

La figura 8 ilustra una realización de la medición del dispositivo del usuario y la gestión del archivo de registro relacionado.

- 55 La figura 9 ilustra una aplicación factible de la clasificación/categorización de contexto y asociaciones relacionadas para su uso en relación con la presente invención.

La figura 10 ilustra una realización del procedimiento de revisión de muestras para propósitos de selección de variables de calibración y categoría (colapso) para permitir la determinación del valor de control y la calibración de datos a nivel de usuario.

La figura 11 es un diagrama de flujo de nivel superior de un procedimiento de acuerdo con la presente invención.

5 **Descripción detallada**

En la sociedad moderna, la fragmentación de dispositivos es uno de los factores que cambian la forma en que las personas se socializan, se comunican, consumen medios, y realizan tareas personales y profesionales. La interacción de los consumidores con los medios y recursos digitales se ha vuelto recientemente móvil: usan teléfonos inteligentes, tabletas, ordenadores portátiles, ordenadores de escritorio, etc. para ejecutar dichas tareas, y muchos incluso utilizan múltiples dispositivos para estos fines diariamente.

El marco descrito en el presente documento se puede realizar como un sistema que se utiliza o es al menos capaz de ser utilizado para medir una gran variedad de características sobre o alrededor del comportamiento de los consumidores expresado a través de una variedad de dispositivos digitales. Las características que se pueden determinar incluyen, pero no se limitan a, el uso por parte de los consumidores, o en general de los usuarios, de varios dispositivos, la base instalada de dichos dispositivos, las acciones del usuario, es decir, lo que los usuarios hacen con sus dispositivos, y por lo tanto, averiguan las formas en que los usuarios utilizan aplicaciones (apps) y sitios web a través de estos dispositivos, cuáles son las aplicaciones, los sitios web o los medios que los usuarios descargan y/o interactúan con los mismos, y cómo estas aplicaciones, servicios y contenido se distribuyen a los dispositivos en primer lugar. Se puede establecer una relación entre la descarga de aplicaciones y su uso real. Se pueden analizar las interacciones económicas relacionadas con o utilizando dispositivos digitales. Se pueden investigar las diferencias y relaciones en el uso en múltiples dispositivos digitales.

Las características anteriores pueden ser inspeccionadas a diferentes niveles de detalle (es decir, en todas las aplicaciones, para varias categorías de aplicaciones, o para aplicaciones específicas), para diferentes audiencias (es decir, grupos de usuarios/consumidores), y/o durante diferentes períodos de tiempo (por ejemplo, de forma diaria, semanal, mensual) según lo previsto por las realizaciones del sistema y el procedimiento de acuerdo con la presente invención.

Mediante el uso de la medición pasiva longitudinal, el marco que se describe en el presente documento es aprovechado para medir el comportamiento digital observando, utilizando ventajosamente muestras en curso a gran escala (por ejemplo, paneles Verto Partner™) para capturar la cola larga de la actividad fragmentada. Los datos obtenidos se pueden calibrar y controlar la calidad sobre la base de una serie de fuentes de datos sustancialmente independientes que incluyen, por ejemplo, datos de estratificación derivados de la muestra para la calibración de la distribución demográfica, un panel longitudinal cuidadosamente estratificado con medición pasiva para perfiles de comportamiento y datos de censos para calibración de comportamiento y escalado.

Un concepto que subyace a las prestaciones del sistema descrito en el presente documento reside en la capacidad para tender conectar siguiendo, por ejemplo, las siguientes tres categorías distintas de información en conjunto:

- 1: datos de distribución de hardware indicativos de la penetración y distribución de diferentes plataformas y dispositivos (móviles), que reflejan los cambios en curso en el mercado de hardware (móvil);
- 2: datos de distribución y monetización de apps (aplicaciones) indicativos de la penetración y la distribución de diferentes aplicaciones (móviles), datos que cuantifican los ingresos generados por las ventas de estas aplicaciones y las compras dentro de la aplicación realizadas por los usuarios; y
- 3: uso y compromiso de la aplicación/contenido: datos que muestran cómo los consumidores interactúan y se involucran con el contenido (móvil) (por ejemplo, aplicaciones, widgets, sitios web).

Además de, o en lugar de, los análisis relacionados con la plataforma móvil/dispositivo/app, el análisis correspondiente relativo a los productos sustancialmente no móviles o no portátiles podría llevarse a cabo siempre que estén disponibles suficientes datos de origen.

La figura 1 ilustra una realización de un sistema de acuerdo con la presente invención desde el punto de vista conceptual y centrado en el dispositivo en 100.

Un sistema 110 electrónico que contiene un número de servidores 108 al menos funcionalmente conectados se proporciona para actividades como adquisición de datos y análisis con respecto al comportamiento en línea de los usuarios 102 (típicamente personas) y otras actividades relativas a la red conectable, preferiblemente Internet 107 conectable, dispositivos 104 de usuario, tal como ordenadores de escritorio, ordenadores portátiles, tabletas, teléfonos celulares y, en particular, teléfonos inteligentes, dispositivos conectados como varios dispositivos 104a, 104b auxiliares (IU de tipo goggle, TV inteligente u otros dispositivos UI, por ejemplo). Dependiendo de la realización y del punto de vista que se tome, también se puede considerar que el sistema 110 incluye varios dispositivos 104 de usuario, opcionalmente dispositivos 104a, 104b auxiliares, y otros elementos, tal como fuentes de datos.

Además de los datos a nivel de usuario capturados utilizando dispositivos 104, 104a, 104b de usuario, preferiblemente a través de una lógica de aplicación de dosificación/recogida de datos 105 (software), que puede ser llamado como medidor de investigación, siempre en el mismo, datos a nivel censo indicativos de la distribución de las características entre toda la población objetivo o, por ejemplo, un subconjunto completo y rigurosamente definido de la misma, se obtiene a partir de varias fuentes de datos independientes del usuario 107a, tal como servicios en línea, que pueden incluir al menos un operador de red, red de publicidad, intercambio de anuncios, plataforma de análisis de sitios o aplicaciones, pasarela de red, proxy de red y/o plataforma(s) de distribución/entrega de contenido, tal como plataformas de administración de contenido o tiendas de aplicaciones. Los datos desde estas fuentes 107a de datos pueden proporcionarse por los dispositivos digitales asociados conectados a la red, tales como servidores, mediante un procedimiento de sondeo o recuperación iniciado por el sistema 110 a través de interfaces de programación de aplicaciones (API) aplicables, y/o mediante un mecanismo de empuje mediante la(s) fuente(s) de datos/servidor(es) 107a de acuerdo con la lógica predeterminada, por ejemplo, en respuesta al cumplimiento de una condición de activación, que puede incorporar una condición basada en temporizador (transmisión de datos diaria, semanal o mensual entre otras opciones). El sistema 110 está configurado para recopilar, clasificar y ensamblar las entidades observadas en los datos obtenidos en una ontología categorizada multinivel por medio de un esquema de metadatos ontológicos y varias técnicas de procesamiento diferentes para servir adecuadamente a las entidades 102a objetivo, que pueden incluir una o más compañías, personas o básicamente cualquier parte o cliente objetivo que solicite o esté interesado en los resultados del análisis. A saber, estas partes recibirán el tipo de entregable(s) deseado(s) que puede(n) incluir, pero no se limita a, informes electrónicos en forma de archivos digitales u otras representaciones, que incluyen de manera ventajosa representaciones visuales, accesibles y/u observables a través de equipos adecuados para el cliente, tales como equipo informático y navegador que se ejecuta en el mismo.

En efecto, la interfaz 112 de usuario para acceder al sistema 110 y/o la inspección de los entregables puede incluir una interfaz de usuario basada en web preferiblemente accesible a través de una aplicación de navegación web común opcionalmente suplementada con módulo(s) de complemento adaptado(s) para mejorar la experiencia de uso. La interfaz basada en la web puede configurarse para permitir al usuario monitorizar convenientemente los datos más recientes o históricos (resultados del análisis), identificar cambios en los datos (tendencias, etc.) y/u opcionalmente realizar análisis ad-hoc utilizando datos históricos o actuales, si es necesario, a través de los componentes interactivos y reactivos de la interfaz de informes basada en la web. Adicional o alternativamente, los entregables pueden proporcionarse a través de una serie de otros canales de salida, por ejemplo, a través de una serie de API accesibles por las aplicaciones y/o servicios de destino, o por correo electrónico.

Como los intereses de las partes pueden variar, el sistema 110 está configurado para determinar preferiblemente de forma dinámica el alcance del análisis y/o del (los) entregable(s) de acuerdo con las necesidades del cliente proporcionados mismo a través de la interfaz 112 de usuario. El alcance puede estar limitado por la población de usuarios, dispositivos usados, aplicaciones usadas, monetización de aplicaciones y/o factor(es) de tiempo, además de otras opciones. Por ejemplo, el sistema 110 puede configurarse para proporcionar una UI 112 que implemente al menos una característica modificable por el usuario para ajustar el alcance del análisis de datos y/o la representación de datos seleccionados del grupo que consiste en: filtrado en términos de, por ejemplo, uno o más grupos objetivo/audiencias, selección de una o más métricas para explorar, selección de un alcance (por ejemplo, categoría de dispositivo, marca del dispositivo, categoría de aplicación, aplicación particular, sitio web, etc.) y período de tiempo deseado (por ejemplo, "enero de 2014") y/o resolución de tiempo (por ejemplo, diaria, semanal, mensual, etc.).

La figura 2 representa, en 200, una realización del sistema desde un punto de vista más conceptual y funcional. El flujo de datos generales (procesamiento) se indica mediante flechas horizontales en la figura entre las actividades de recopilación de datos y las actividades de producción/salida.

En 202, un estudio 208 estructural se indica como una fuente de información para ser utilizada como una base para la calibración de datos en el sistema, por ejemplo.

Preferiblemente, el estudio 208 estructural es o incorpora una encuesta independiente para proporcionar una base independiente, potencialmente junto con los datos 211 del censo, para un número de acciones tales como la determinación de la estratificación de panel deseada (es decir, cuotas) para la contratación del (primer) panel 216 de investigación, y la calibración de los datos resultantes recopilados durante el proceso 202 de recopilación de datos y recopilados o reunidos durante el proceso 204 de producción de datos. El estudio 208 estructural se puede realizar como un estudio fuera de línea para maximizar su cobertura/representatividad, por ejemplo.

El estudio 208 estructural está diseñado para delinear los supuestos estadísticos básicos que describen la población investigada. El estudio 208 estructural se puede utilizar para responder (conceptualmente), por ejemplo, las siguientes preguntas:

- ¿Cuáles son las características de propiedad del dispositivo de la población (adulta) del mercado deseado (por ejemplo, cierto país, estado, región u otra área geográfica)?
- ¿Cuáles son las características sociodemográficas de los propietarios de dispositivos medidos?



El estudio 208 estructural podría ser implementado por la parte que también implementará el sistema descrito en esta invención, o, alternativamente, en ciertas implementaciones de la metodología, el estudio 208 puede ser licenciado o comisionado de un proveedor de estudios de mercado con reputación y creíble, u otro investigador social aplicable. Existen una variedad de procedimientos de investigación que pueden proporcionar datos suficientes para satisfacer los requisitos enumerados anteriormente, y que se pueden licenciar fácilmente o encargar a proveedores de investigación de terceros.

Los datos 211 del censo pueden incorporar datos del censo de red 210, que se pueden obtener a través de tráfico de red o monitorización de la actividad de red, por ejemplo, y específicamente, por ejemplo, almacenar datos 212 de aplicaciones. Preferiblemente, los datos a nivel de censo describen una variedad de temas, que en algunas realizaciones pueden incluir dispositivos, usuarios, servicios, aplicaciones o sitios web. Por lo tanto, cada conjunto de datos incluido en los datos del censo deberá proporcionar características agregadas o a nivel de caso e información sobre la población total a la que se aplica ese conjunto de datos particular. Los datos 211 del censo pueden usarse con los datos 208 estructurales para calibrar los datos recopilados en el proceso 202 de recopilación de datos y los datos recopilados o reunidos en el proceso 204 de producción de datos.

Los metadatos 214 incorporan, o definen, un conjunto de datos ontológicos que se utilizan para definir taxonomías de entidad que sirven para categorizar entidades (por ejemplo, dispositivos móviles, sistemas operativos, sitios web visitados, aplicaciones, etc.) descubiertos u observados durante el proceso 202 de recopilación de datos y el proceso 204 de producción de datos. Estos datos pueden ser utilizados para:

- agrupar elementos de datos relacionados en categorías según sea necesario durante el proceso 204 de producción de datos, y/u
- obtener datos organizados categóricamente en el(los) entregable(s) 206.

El medidor 217 de investigación se puede realizar como software que se obtiene, opcionalmente descargado a, e instalado en dispositivos digitales (por ejemplo, ordenadores, teléfonos inteligentes, tabletas, etc.) para monitorizar y registrar la actividad del mismo. Básicamente, puede actuar como una fuente primaria de datos para el sistema y el marco descrito en esta invención, y sus datos se utilizarán para recopilar datos.

- datos de observación sobre el comportamiento/actividad de los dispositivos monitorizados en el primer panel 216 de investigación (ver más abajo), y/o
- datos de observación sobre el comportamiento/actividad de los dispositivos monitorizados en el segundo panel 218 de investigación, por ejemplo, el Panel Verto Partner™ mencionado anteriormente.

Ahora, finalmente, teniendo en cuenta los paneles 216, 218 en más detalle, el primer panel 216 de investigación es preferiblemente un panel pasivamente dosificado de los usuarios de dispositivos digitales, que podrán ser seleccionados preferiblemente usando un procedimiento seleccionado y definido. En algunas realizaciones, este procedimiento puede involucrar técnicas tales como CAWI (entrevista web asistida por ordenador) facilitada a través de intercepciones emergentes, invitaciones a paneles de acceso en línea, intercepciones de registro oportunistas, CATI (entrevista telefónica asistida por ordenador), CAPI (entrevista personal asistida por ordenador), u otros procedimientos de investigación comúnmente utilizados y bien establecidos.

La composición del panel se determinará ventajosamente en función de las cuotas específicas derivadas del estudio estructural. Los panelistas tendrán el medidor 217 de investigación instalado en sus dispositivos digitales, y el medidor 217 de investigación posteriormente monitorizará pasivamente su comportamiento. La medición pasiva generalmente se refiere a la tecnología en la que el software que realiza el seguimiento de las actividades del usuario o los parámetros contextuales, al menos en una medida significativa, se ejecuta en segundo plano en dichos dispositivos, sin la necesidad de que los usuarios interactúen con ese software por separado. En la instalación de estos medidores de software en dispositivos digitales, los usuarios pueden necesitar completar las etapas relacionadas con la instalación. Un medidor de software que se ejecuta en un dispositivo de este tipo, en algunas realizaciones, también puede proporcionar datos de medición en una serie de otros dispositivos digitales, por ejemplo, los que están conectados funcionalmente a los dispositivos maestros y, por lo tanto, se comunican con los mismos de forma inalámbrica.

Un objetivo del primer panel de investigación 216 es proporcionar una fuente de datos de alta calidad (es decir, altamente controlada, finamente calibrada) para el comportamiento del consumidor. El tamaño de la muestra de investigación proporcionará una profundidad de datos limitada (es decir, disponibilidad limitada de datos de cola larga), sin embargo, el panel proporcionará específicamente datos de alta calidad para esos temas (por ejemplo, aplicaciones, sitios web, fabricantes de dispositivos, grupos demográficos, etc.), cuya audiencia/usuarios es suficientemente grande, si se implementa con cuidado.

En detalle y dependiendo de la realización del medidor 217 de investigación y tras la finalización del proceso 204 de producción de datos, el panel 216 puede estar diseñado para proporcionar datos de alta calidad que responden, por ejemplo, a las siguientes preguntas:

- ¿Cuál es el uso de dispositivos digitales y las tasas de penetración entre los consumidores de un determinado

mercado?

- ¿Con qué recursos (por ejemplo, aplicaciones, sitios web, medios) interactúan los consumidores con sus dispositivos digitales?
- ¿Cómo interactúan los consumidores con los principales recursos (por ejemplo, aplicaciones, sitios web, medios) a través de sus dispositivos digitales?
- ¿Para qué, cuándo, cómo y en qué cantidades los consumidores gastan dinero en dispositivos digitales?
- ¿Cómo varían las preguntas anteriores según los dispositivos al considerar a consumidores con múltiples dispositivos digitales?
- ¿Cuáles son las tendencias en la distribución/uso, uso, reemplazo y/o uso incremental de la aplicación/contenido?

Los datos desde el panel 216 pueden formar o ser usados para formar un entregable principal 206, y se conforman a la forma de los datos de salida derivados en base al segundo panel de investigación 218 (por ejemplo, Panel Verto Partner™).

Como se ha mencionado en el presente documento, el panel 216 pueden ser reclutado sobre la base de un conjunto de cuotas estratificadas y la estratificación del panel puede determinarse basándose en los resultados del estudio 208 estructural.

Antes de ser invitado a participar en el primer panel 216 de investigación (es decir, antes de ingresar al marco de muestreo), el encuestado debe cumplir con una serie de criterios básicos, que pueden incluir diferentes requisitos con respecto a, por ejemplo, la edad y residencia del encuestado, la propiedad del dispositivo/red o accesibilidad, historial de participación en encuestas, etc.

Los panelistas que participan en el panel 216 pueden ser compensados de una manera deseada. El mecanismo de gratificación y el procesamiento de incentivos para cada panelista en particular se determinará en colaboración con el proveedor de la muestra que originó dicho panelista. El incentivo puede ser digital y asignarse digitalmente, tal como acceso o derechos de acceso/usuario elevados a contenido digital, servicios, aplicaciones, etc.

El segundo panel 18 de investigación, por ejemplo, Panel Verto Partner™, es preferiblemente un panel medido de forma pasiva de los usuarios de dispositivos digitales reclutados más ventajosamente por medio de muestreo de conveniencia. En algunos casos, el reclutamiento puede ser indirecto considerando, por ejemplo, escenarios en los que el medidor 217 de investigación se incluye como incrustado en ciertas aplicaciones para usuarios finales que los usuarios descargan principalmente para fines distintos a participar en el estudio de investigación. El panel constará de o contendrá usuarios que hayan:

- instalado el medidor 217 de investigación incluido con una o más aplicaciones móviles;
- proporcionado al sistema sus datos de perfil demográfico; y/o
- acordado proporcionar datos anónimos al sistema.

Debido a que la composición del panel es, de forma predeterminada, sustancialmente incontrolada (es decir, "básicamente todos son bienvenidos") o casi incontrolada, la función del panel 218 es desarrollar una muestra extremadamente grande capaz de proporcionar un nivel profundo de datos de cola larga. El panel 218 puede configurarse específicamente para proporcionar datos complementarios para una gran cantidad de temas (por ejemplo, aplicaciones, sitios web, dispositivos, etc.).

Y como la representatividad demográfica y de comportamiento de la muestra es incontrolada en vista de lo anterior, los datos de los mismos no serán utilizados para calibrar otros datos. En cambio, los datos del panel 218 se calibran por sí mismos durante el proceso 204 de producción de datos con datos que son más representativos y controlados por la naturaleza, como los datos del primer panel 216 de investigación y otros datos recopilados en el proceso 202 de recopilación de datos.

En cuanto a las características técnicas y la filosofía relacionadas del panel 218 en más detalle, debido a que los temas sobre los que el marco que se describe en el presente documento está destinado a informar son tan fragmentados, en circunstancias prácticas, es imposible proporcionar datos completos/en profundidad en comportamiento digital sin un tamaño de muestra extremadamente grande. Reclutar una muestra extremadamente grande utilizando los procedimientos rigurosos diseñados para el primer panel 216 de investigación (ver arriba) sería prohibitivamente costoso en base a cada panelista. Por lo tanto, el primer panel 216 se complementará a través del desarrollo del segundo panel 218, preferiblemente reclutado en una base de conveniencia (no controlada o "captura como se pueda capturar").

El panel 218 preferiblemente no es de tamaño limitado. Sin embargo, en algunas realizaciones, los datos del panel 218 solo pueden incluirse en el(los) entregable(s) 206 cuando el número de panelistas seleccionados mediante un control de calidad aplicado y un proceso de validación excede un número predeterminado, por ejemplo, 20000 (veinte mil).

En realizaciones preferidas, la composición del segundo panel 218 no se controlará a priori, pero los estados de

validación están diseñados ventajosamente para subseleccionar los encuestados pertinentes para su inclusión como panelistas validados (en pestaña) post factum.

5 Los panelistas del segundo panel 218 pueden ser reclutados de una población de usuarios de dispositivos móviles (es decir, usuarios de iOS, Blackberry, Android, y/o dispositivos Windows Mobile, por ejemplo). Por ejemplo, la fuente específica de los panelistas pueden ser aquellos usuarios que han instalado una o más aplicaciones móviles que están integradas con el medidor de investigación 217, o, por ejemplo, SDK de investigación (kit de desarrollo de software).

10 Todo usuario que haya instalado una aplicación integrada puede, en principio, ser invitado a participar en el segundo panel 218. La invitación se puede proporcionar a través de la aplicación, por ejemplo, mediante un mensaje visual. Esta participación se basa en la aceptación voluntaria del usuario. Sin embargo, para requisitos normativos y éticos, puede ser necesario que, junto con la inclusión voluntaria, el usuario indique su año de nacimiento como una característica de calificación previa (por ejemplo, los usuarios menores de 18 años de edad se excluirán automáticamente).

15 Se puede realizar una encuesta de reclutamiento para recopilar datos básicos/mínimos sobre el encuestado para ponderar adecuadamente sus datos. La información demográfica básica recopilada en la encuesta de reclutamiento del panel 218 podría incluir, por ejemplo, tipo de dispositivo, operador móvil, género, educación, ubicación geográfica, tamaño del hogar, etc.

20 Esta información a los datos de ponderación de panelista mega puede derivarse técnicamente a partir de un número de fuentes que utilizan opcionalmente de forma automática, por ejemplo, las observaciones basadas en software, o integran terceras partes seleccionadas para describir los panelistas, o usando alguna forma de ventanas emergentes o cuestionarios para actualizar esto para los panelistas.

25 Los datos 206 entregables son la salida final del sistema descrito en el presente documento, pero, obviamente, el cliente 102a que los recibe puede cultivar aún más, por ejemplo, con herramientas propietarias, etc., si surge una necesidad. El entregable 206 contiene información calibrada, procesada y categorizada, por ejemplo, sobre el uso del dispositivo, el uso de la aplicación/servicio y el comportamiento del usuario, que se pondrá a disposición de los clientes de forma opcional o parcial a través de una interfaz de usuario basada en la web. Alternativamente, por ejemplo, podría utilizarse una aplicación nativa para consumir la información.

30 El(los) entregable(s) 206 se puede(n) realizar como una sola base de datos, unificada en el extremo posterior, por ejemplo. Sin embargo, a cada cliente se le pueden asignar derechos de usuario específicos para acceder a porciones limitadas de los mismos a través de las interfaces de usuario disponibles, tal como una interfaz basada en la web.

Por ejemplo, el(los) entregable(s) puede(n) incluir al menos un elemento seleccionado del grupo que consiste en: distribución de hardware, distribución de contenido, uso de app/contenido y compromiso, y análisis de aplicaciones.

35 La figura 11 muestra un diagrama 1100 de flujo genérico de una realización de un procedimiento de acuerdo con la presente invención que implica los temas y las características explicadas anteriormente.

40 En el arranque 1102, el sistema se intensifica, que puede referirse a la adquisición y a la configuración de hardware y software relacionado. El sistema puede incluir varios dispositivos de servidor, por ejemplo, dispuestos en un entorno informático en la nube, por ejemplo. Sin embargo, el reclutamiento "silencioso" de panelistas para al menos el segundo panel de usuario ya se ha iniciado en esta etapa, por ejemplo, a través de muestreos por conveniencia, y el software de medición, registro e informe relacionado se entrega a una multitud de dispositivos de usuario, tal como terminales móviles. u otros terminales, preferiblemente como incluido con otro software. Aun así, por ejemplo, se pueden seleccionar los proveedores de datos a nivel de censo y establecer los canales de comunicación de datos necesarios.

45 En 1104, se ejecuta el estudio estructural para determinar la estratificación (cuotas) del primer panel de investigación, las propiedades del universo medido (datos demográficos, etc.), que se utiliza para la calibración de datos de panel, etc. El reclutamiento de los panelistas, especialmente para el primer panel 216, puede ejecutarse en consecuencia.

En 1106, se obtienen los datos a nivel de censo preferiblemente de una pluralidad de fuentes, tales como tiendas de aplicaciones, proveedores de datos de red, etc.

50 En 1108, los datos a nivel de usuario se reciben de los panelistas.

En 1110, el procesamiento de datos se realiza incluyendo la organización de los datos. La calibración de los datos a nivel de usuario, por ejemplo, con datos de estratificación derivados de la muestra (a partir del estudio 208 estructural) y/o datos a nivel de censo, se realiza como se describe con más detalle a continuación.

En 1112, se construye(n) el(los) entregable(s) basado(s) en los datos de censos y de nivel de usuario recogidos. El

elemento 1116 se refiere a la entrada del cliente que da forma al contenido y/o la forma de los entregables (el sistema puede recibir y tener en cuenta dicha entrada en varias etapas del proceso, aunque no se indica en la figura hasta el elemento 1112).

En 1114, la ejecución del procedimiento se acaba.

5 Como se indica en la figura mediante las diversas flechas de bucle de puntos, la ejecución de elementos puede ser, y normalmente es, repetitiva y/o iterativa para ser clara para una persona experta en la técnica sobre la base de esta descripción. Además, el orden de ejecución puede variar dependiendo, por ejemplo, de la disponibilidad de nuevos datos y también el procesamiento paralelo o alternativo de elementos puede tener lugar, y típicamente tendrá lugar en implementaciones de la vida real considerando, por ejemplo, una calibración iterativa potencial para los datos a nivel de usuario.

10 Con referencia a la figura 3, en 300, se representa una realización de enfoque de múltiples dispositivos factible para ser utilizado en conexión con la presente invención. En general, la comunicación entre el usuario 102 y sus dispositivos 104, 104a se ilustra con líneas continuas y la comunicación entre dispositivos se indica con líneas discontinuas. De hecho, en realizaciones preferidas de la presente invención, se pueden rastrear múltiples dispositivos de un solo usuario, incluso si uno o más de esos dispositivos no puede, o no estará, configurado para ejecutar al menos tecnologías de recogida de datos basadas en software o hardware a gran escala por razones prácticas. En particular, el sistema se configura preferiblemente de modo que, aunque los usuarios 102 interactúen con dispositivos conectables o portátiles o interfaces de usuario externas (por ejemplo, dispositivos con reloj inteligente o gafas digitales de pantallas digitales), se pueden recopilar los datos de medición relacionados. En estos casos, el sistema utiliza los llamados dispositivos maestros 104 entre medio, a los que se conectan dichos dispositivos conectados o interfaces de usuario 104a, para realizar una o más de las siguientes funciones:

1. Rastrear el tráfico, las señales u otros eventos que tienen lugar entre el(los) dispositivo(s) 104a conectado(s) y el dispositivo 104 maestro.
- 25 2. Realizar una encuesta pasiva (por ejemplo, en función de algunos activadores de eventos contextuales) o activa (en función de intervalos predefinidos, por ejemplo) de las interfaces de programación de aplicaciones (API) de los dispositivos 104a conectados para recopilar datos de medición.
3. Almacenar los datos de medición recopilados en el dispositivo maestro temporalmente o durante un período de tiempo predeterminado.
4. Enviar los datos al servidor 108 de recopilación de datos basado en la red del sistema para fines de análisis.

30 Preferiblemente, el dispositivo maestro está configurado para realizar el seguimiento, almacenar y enviar los datos de tal manera que los datos originalmente en relación con un dispositivo particular (ya sea el propio dispositivo maestro o cualquier dispositivo conectado) más tarde se sigue de vuelta al mismo dispositivo. En una implementación factible, si bien el usuario puede tener asignado un determinado código de identificación de usuario o número (ID), cada uno de los dispositivos, no importa si son maestros 104 o dispositivos 104a conectados (esclavos), tienen su propio número de identificación único del dispositivo también, aunque algunos de ellos pueden estar conectados entre sí.

Este tipo de implementación que permitirá romper los datos de uso obtenidos abajo en diferentes dispositivos e interfaces de usuario entre otros beneficios potenciales durante la fase de análisis. Además, en ciertas implementaciones, el motor que maneja el seguimiento/mediciones, que puede residir en un dispositivo maestro o en un servidor remoto, además de asignar, por ejemplo, números de ID de usuario y/o números de ID de dispositivo (correspondientes a los dispositivos maestros), puede estar configurado para asignar o registrar ID de interfaz de usuario a interfaces de usuario/dispositivos conectados que están conectados al dispositivo 104 maestro.

40 De hecho, en la puesta en práctica donde las mediciones se ejecutan en un grado significativo y dirigido por el dispositivo 104 maestro (por ejemplo, que el dispositivo se utiliza como el almacenamiento de datos y una pasarela de entrada hacia la red, y para manejar la identificación/registro de las diferentes interfaces de usuario que se están usando), pero las acciones de interacción y de usuario pueden tener lugar a través de un dispositivo 104a esclavo, unido y conectado, por ejemplo, se pueden realizar las siguientes etapas técnicas:

1. El dispositivo maestro registra las interfaces de usuario nuevas o existentes que están conectadas al dispositivo maestro, ya sea a través de la escucha dinámica de nuevas conexiones, o alternativamente, mediante el sondeo de información periódica sobre dispositivos emparejados o conectados, por ejemplo;
- 50 2. La información de estado, que incluye preferiblemente un latido cardíaco (el dispositivo está encendido, incluso si no se produce un acoplamiento activo con el contenido, los servicios, las aplicaciones), se realiza un seguimiento de cada dispositivo.

55 En el sistema descrito en el presente documento para realizar un seguimiento de múltiples dispositivos o comportamientos de múltiples UI, el dispositivo 104 maestro deberá registrar ventajosamente eventos y otros datos de medición, potencialmente incluyendo capturas de pantalla, de forma centralizada, sin tener que sobrecargar el dispositivo 104a conectado con estas actividades. El dispositivo 104 maestro puede configurarse para usar las API disponibles ampliamente utilizadas y otras llamadas de función para recopilar estos datos, o

alternativamente/adicionalmente en algunas implementaciones podría rastrear canales de tráfico de señalización más fundamentales, datos HTTP (tal como el dispositivo principal que se usa como tipo de pasarela de enlace de red), u otra API o API de nivel inferior.

5 Con referencia a la figura 4, las partes internas de una realización de un sistema 400 de acuerdo con la presente invención con funcional/lógica 401A y puntos 401B de vista de hardware se muestran en la misma.

El elemento 412 se refiere a la lógica de control principal que se ocupa de la adquisición de datos de diversas fuentes, la transferencia de datos entre diferentes entidades, la gestión de la configuración, etc.

El elemento 410 se refiere a una serie de depósitos de datos tales como bases de datos para almacenar los datos recopilados y procesados.

10 El elemento 414 se refiere a esquemas de metadatos ontológicos utilizados para organizar y clasificar los datos en un conjunto de datos multivalente.

El elemento 416 se refiere a la interfaz de usuario para generar el(los) producto(s) que contienen los datos de múltiples niveles integrados y categorizados obtenidos al procesar y analizar los datos a nivel de usuario recopilados y los datos a nivel de censo.

15 El elemento 408 se refiere a la lógica de recopilación de datos a nivel de usuario en el extremo del sistema o servidor, mientras que el elemento 402 se refiere a un medidor de investigación 217 o (SDK de investigación), es decir, la lógica se ocupa de la medición de datos, el registro, el procesamiento potencial, y el reenvío en dispositivos de usuario y, por ejemplo, en conexión con servicios en línea/móviles. Preferiblemente, la lógica se prepara de tal manera que se puede conectar fácilmente a cualquier servicio en línea o móvil, o aplicación, y como la gente usa esos servicios o aplicaciones particulares, la lógica se activa en segundo plano, ejecuta las mediciones y puede enviar datos de medición al servidor(es) del sistema sugerido. Todo esto puede producirse sin necesariamente proporcionar ninguna interfaz o posibilidad de configuración a la "aplicación hospedadora" para afectar al funcionamiento de la lógica. Sin embargo, en algunas implementaciones, puede haber una serie de "bucles de retroalimentación" seleccionados integrados en el módulo lógico, de modo que pueda proporcionar señales, datos u otra cosa, de vuelta a la aplicación hospedadora.

20

25

Una lista de aplicaciones integradas con la lógica 402 se mantiene preferiblemente en todo momento para fines de supervisión/gestión. A cada usuario (móvil) que instala una aplicación integrada se le puede dar la oportunidad de participar (optar por) el segundo panel 218 de investigación. Esta población de usuarios que han instalado una aplicación integrada puede establecer así el marco de muestra para el segundo panel 218.

30 El elemento 404 se refiere de manera correspondiente a la lógica de recopilación de datos a nivel de censo. Aunque no se indica explícitamente en la figura, también en este caso, el SDK de investigación o la entidad de recopilación de datos correspondiente instalada en una entidad remota, por ejemplo, en los servidores que ejecutan y/o alojan un servicio en línea, pueden aplicarse para fines de recopilación de datos.

35 El elemento 406 se refiere a la lógica de análisis de datos configurada para verificar, combinar, calibrar, procesar y organizar los datos recopilados utilizando, por ejemplo, el(los) esquema(s) 414 para obtener el tipo deseado de entregables para la salida a través de la UI 416.

40 En cuanto al hardware 401B, cada uno de los dispositivos aplicados, tales como servidor(es) y dispositivo(s) de usuario (dispositivos maestros y dispositivos conectados) puede incluir varios dispositivos 420 de procesamiento, tal como microprocesadores, microcontroladores, procesadores de señales, etc. para el procesamiento de instrucciones y datos. Además, pueden contener memoria (chips) 422 para almacenar las instrucciones del programa y otros datos tales como los datos recopilados a nivel de usuario y/o nivel de censo, datos procesados, esquema(s) ontológico(s), entregables, etc.

45 La interfaz de datos 426 puede referirse a una interfaz de interfaz de punto a punto o de red que puede ser por cable o inalámbrica, por ejemplo. Un dispositivo de tipo servidor puede contener típicamente un adaptador de red (cableado) con, por ejemplo, Ethernet u otros medios de conectividad de tipo LAN, mientras que un dispositivo de usuario móvil puede contener un número de transceptores inalámbricos, por ejemplo, un transceptor celular (por ejemplo, 3G) y/o WLAN (red de área local inalámbrica o transceptor Bluetooth, para acceder a Internet y/o comunicarse con otros dispositivos cercanos. También se puede incluir un lector de etiquetas (RFID (identificación por radiofrecuencia) o NFC (comunicación de campo cercano), por ejemplo. La interfaz de usuario puede comprender una pantalla 424 y un dispositivo 430 de entrada de datos, tal como un teclado, un teclado, una interfaz táctil (pantalla táctil, por ejemplo), etc. Sin embargo, los dispositivos generalmente contienen varios otros elementos 428, tal como uno o más sensores, componentes de la fuente de alimentación, carcasa, etc.

50

55 El código de programa (software) para llevar a cabo las etapas del procedimiento y, en general los procedimientos descritos en este documento por un número de dispositivos informáticos, puede proporcionarse generalmente en un medio de soporte tal como tarjeta de memoria o disco óptico, o transferirse a través de un medio de comunicación por cable o inalámbrico.

- Con referencia a la figura 5, las relaciones entre diferentes poblaciones potenciales medidas a través de los datos 211 del censo se muestran en 500. A saber, los datos que se incluyen en los datos 211 del censo pueden subdividirse en conjuntos de datos independientes y no relacionados, cada uno de los cuales informa de datos sobre un subconjunto particular de la población general de usuarios (ya sea "usuarios de ordenadores" 510, "usuarios de web móvil" 512 , o "usuarios móviles no web" 508).
- Es probable que las tres poblaciones mencionadas se solapen: un individuo puede ser simultáneamente un usuario de ordenador, un usuario web de dispositivo móvil, y un usuario de la aplicación del dispositivo móvil. Sin embargo, los diferentes proveedores de datos del censo generalmente tendrán un alcance diferente entre cada una de estas tres subpoblaciones.
- Para que los datos 211 del censo proporcionen datos fiables para su uso en la calibración, los proveedores de datos del censo se seleccionan preferentemente de manera que al mismo tiempo se maximice el alcance entre cada una de estas tres poblaciones (preferiblemente cerca del 100 %), y se minimiza la probabilidad de que un usuario particular se cuente dos veces en todos los proveedores de datos del censo (preferiblemente cerca del 0 %).
- En la práctica, una manera de acercarse de manera creíble al 100 % del alcance en las tres subpoblaciones medidas es combinar múltiples fuentes de datos del censo. Sin embargo, debido a que es necesario minimizar la probabilidad de que un usuario único particular haya sido contado dos veces, la superposición de audiencia entre estas fuentes de datos del censo debe minimizarse a priori. Se recomienda que uno que implemente el sistema de la presente invención seleccione cuidadosamente a los proveedores de datos del censo para determinar con precisión su alcance y superposición en cada subpoblación medida.
- Los datos del censo se procesan y se comparan preferiblemente de manera oportuna (por ejemplo, mensualmente) con otras fuentes de datos, con fines de calibración, y para ajustar el cálculo de corrección de modo que las estimaciones publicadas, por ejemplo, sobre el número de dispositivos en uso, o los nuevos dispositivos activados, estén lo más cerca posible de los números verdaderos.
- Los datos del censo pueden obtenerse en un formato de nivel sin procesar, que incluye, por ejemplo, una tabla que divulga, para cada campo que se comunica sobre el tipo de usuario o dispositivo, el número nominal o relativo de usuarios únicos vistos durante un cierto período de tiempo, y/o el número nominal o relativo de acciones observadas durante un período de tiempo determinado, por ejemplo, para un país determinado.
- En algunas realizaciones, el sistema puede estar configurado para utilizar cuerdas de nivel en bruto en la información de nivel en bruto suministrada para identificar el tipo de usuario, el tipo de dispositivo, y la aplicación utilizada, de forma automática o semiautomática (por ejemplo, con curación humana). Las cadenas de nivel en bruto pueden incluir, por ejemplo, piezas de datos alfanuméricos que sirven como identificadores. Del mismo modo, en función de, por ejemplo, cualquier ID relacionado con el dispositivo, tal como números de serie, temporales o permanentes, se puede contar un número único de dispositivos para cualquier dispositivo y/o parte específica de la aplicación de los datos del censo.
- En algunas realizaciones de esta invención, los paneles 216, 218 pueden ser reclutados sin medios técnicos de control de polarización de comportamiento en sus respectivas estratificaciones de muestra.
- Aunque el estudio 208 estructural proporciona alguna medida de control demográfico y calibración, su base declarativa hace que su valor de comportamiento sea un tanto sin sentido para la calibración de la representatividad de comportamiento de los paneles 216, 218.
- En algunos casos, los datos del censo pueden contener determinados puntos de datos que reflejan el comportamiento de una subpoblación entera definida de usuarios, dispositivos o servicios. Dichos datos del censo se pueden usar para calibrar cualquier sesgo de comportamiento involuntario que resulte de la incapacidad mencionada anteriormente para incorporar características de comportamiento en la estratificación de la muestra de los paneles 216, 218 respectivos. Un ejemplo de datos de censo que pueden proporcionar datos de comportamiento a nivel de censo para su uso en calibración es la información a nivel de censo sobre las aplicaciones disponibles en las tiendas de aplicaciones móviles (tiendas de aplicaciones), que se representan como datos 212 de la tienda de aplicaciones.
- Cambiando a la figura 6, los datos de la tienda de aplicaciones o datos similares pueden ser recopilados inteligentemente por robot(s)/rastreador(es) programado(s) para escanear interfaces públicas o de propiedad pública, por ejemplo, periódicamente, cuya realización se ilustra en la figura en 600.
- Por ejemplo, diariamente (una vez cada 24 horas), el sistema de la presente invención puede configurarse para conectarse a las tiendas de aplicaciones monitorizadas u otras plataformas de suministro/distribución de contenidos más o menos correspondientes, opcionalmente a través de API adecuados, e iterar a través de cada aplicación enumerada en la tienda de aplicaciones. Con respecto a cada aplicación, puede configurarse para registrar, por ejemplo, uno o más puntos de datos como la tienda de aplicaciones en cuestión, una marca de tiempo, el nombre de la aplicación en cuestión, etc.

Con respecto a cada aplicación, puede configurarse para registrar, por ejemplo, uno o más de los siguientes elementos de datos:

- Tienda de aplicaciones. Un identificador único que indica la tienda de aplicaciones desde la que se recopilaron los datos.
- 5 • Marca de tiempo. Una marca de tiempo UTC indicativa de cuándo se recopilaron los datos de la tienda de aplicaciones.
- Identidad de la aplicación. El código de ID o nombre (u otro identificador único) de la aplicación en cuestión.
- Información sobre la posición de la aplicación en la(s) lista(s) de clasificación.
- Categoría de la aplicación. Datos que describen cómo se clasifica la aplicación.

10 El rastreador de la tienda de aplicaciones puede tener una unidad 608 de servidor maestro que es capaz de almacenar datos en una base 612 de datos en bruto, y el acceso al almacenamiento 610 de metadatos, que contiene información sobre las aplicaciones ya escaneadas y categorizadas y, por ejemplo, el estado del rastreador de la tienda de aplicaciones.

15 Entonces, el rastreador de la tienda de aplicaciones puede configurarse para utilizar varias cuentas de usuario diferentes y/u otras características orientadas a emuladores para "simular" ser o emular un dispositivo inalámbrico particular o un usuario particular, por ejemplo, y/o usando ya sea mediante proxis/túneles VPN 606, o accediendo directamente a las API de la tienda de aplicaciones (a veces públicas, a veces privadas), solicitar información ya sea en parches o de una en una, desde las diferentes instancias de la tienda 602, 604 de aplicaciones, por ejemplo, en diferentes países, y con ello datos relevantes a la plataforma. Por lo tanto, el sistema puede diseñarse para usar características orientadas a emuladores y/u otras tecnologías, tal como la creación de túneles y esquemas de direcciones IP dinámicas, para rastrear efectivamente las API de la tienda de aplicaciones de diferentes proveedores, países y localizaciones.

25 Debido a que ninguna, o al menos no todas, las tiendas de aplicaciones disponibles generalmente publican (públicamente) el número real de descargas para cada aplicación, el número de descargas puede estimarse en función de una variedad de puntos de datos recopilados para cada aplicación, por ejemplo, pero no limitado a clasificaciones de aplicaciones específicas de cada país, puntuaciones de calificaciones de usuarios, etc.

30 La figura 7 representa, en 700, diferentes aspectos de una realización con respecto a una nueva técnica de estimación de estadísticas de descarga, por ejemplo, número de descargas, en base a la información disponible en forma de, por ejemplo, clasificaciones y/o calificaciones (con respecto a las entidades objetivo de la estimación). Con esta solución, es posible, entre otros usos, producir estimaciones para un conjunto completo de aplicaciones.

35 Muchas tiendas de aplicaciones, como iTunes™ y Google Play™, publican algún tipo de listas principales y, por ejemplo, clasificaciones para aplicaciones de teléfonos inteligentes, pero no publican, al menos de manera integral, los números reales de descargas de diferentes aplicaciones. Sin embargo, la información de descarga sería de gran interés para muchos jugadores en el mercado de aplicaciones móviles, desde desarrolladores de aplicaciones hasta investigadores de mercado. Preferiblemente, todavía hay información de clasificación disponible públicamente sobre las aplicaciones en varias fuentes, tal como las tiendas de aplicaciones mencionadas anteriormente. Para estimar la cantidad de, por ejemplo, descargas diarias para las aplicaciones enumeradas en una tienda de aplicaciones utilizando información disponible públicamente sobre cada aplicación y, entre otros factores, la posición de clasificación publicada para cada aplicación por la(s) tienda(s) de aplicaciones en cuestión, se describen a 40 continuación una serie de enfoques.

45 Antes de continuar con los detalles adicionales, en primer lugar, se supone que existe una relación matemática bien definida entre la posición de clasificación en una tienda de aplicaciones determinada y el número de descargas (diarias) realmente experimentadas por una aplicación determinada, y que los datos de posición de clasificación están disponibles. Sin embargo, se requiere un conjunto de datos verdaderos básicos para determinar una serie de parámetros del modelo.

50 El uso de un modelo matemático para estimar descargas por referencia de tiempo deseada, tal como las descargas diarias mencionadas anteriormente tomadas aquí como ejemplo, se habilita por los datos de posición de clasificación (diarios) obtenidos, por ejemplo, de la(s) tienda(s) de aplicaciones para un determinado conjunto de aplicaciones y una serie de parámetros de modelo determinados correctamente derivados, por ejemplo, desde una fuente independiente de datos verdaderos básicos.

Con respecto al propio modelo, se ha sugerido un modelo simplista de ley de energía, pero hay una serie de otras opciones que brindan diversos grados de precisión cuando se comparan con los datos verdaderos básicos. En particular, se exponen los siguientes modelos:

*Exponencial:*

55 
$$d(r) = a * \exp(b * r)$$

*Log-normal:*

$$d(r) = a/r^b * \exp(c * \log^2(r))$$

*Ley de potencia con corte exponencial:*

$$d(r) = a/r^b * \exp(c * r)$$

*Lomax:*

5

$$d(r) = a(1 + b * r)^c$$

en el que  $d(r)$  se refiere a descargas (diarias) y  $r$  a la clasificación. Aunque el modelo exponencial propuesto tiene solo dos parámetros, los modelos restantes tienen tres parámetros que permiten el modelado de relaciones más complejas. Por ejemplo, el modelo log-normal incluye los siguientes parámetros: parámetro de escala  $a$ , y parámetros de forma  $b$  y  $c$  que determinan la inclinación y la curvatura del modelo.

10

El procedimiento propuesto implica las siguientes etapas:

1. Supongamos un modelo 702. En la primera etapa del procedimiento propuesto, se asume el modelo matemático que debe aplicarse. Por ejemplo, los modelos *exponenciales*, *log-normal*, *ley de potencia con corte exponencial*, y *lomax* pueden considerarse (y preferiblemente se consideran).

15

2. Ajustar los parámetros 704. Dado el modelo asumido, se usan los datos 712, 714 externos como la posición de clasificación conocida y se descargan los datos para ajustar los parámetros al modelo para obtener el modelo 706 actualizado. Este proceso se puede facilitar al convertir los modelos en sus versiones logarítmicas:

*Ley de potencia:*

$$\log(d(r)) = \log(a) + b * \log(r)$$

*Exponencial:*

20

$$\log(d(r)) = \log(a) + b * r$$

*Log-normal:*

$$\log(d(r)) = \log(a) - b * \log(r) + c * \log^2(r)$$

*Ley de potencia con corte exponencial:*

$$\log(d(r)) = \log(a) - b * \log(r) + c * r$$

25

*Lomax:*

$$\log(d(r)) = \log(a) + c * \log(1 + b * r)$$

3. Para cada aplicación de destino, se construye la estimación 710 de descargas utilizando 708 el modelo actualizado. Una vez que haya ajustado los parámetros al modelo, se aplica, por ejemplo, uno de los dos algoritmos descritos a continuación para estimar el número de descargas para una aplicación particular, dependiendo de los datos 716 externos disponibles.

30

El algoritmo 1 se refiere a situaciones en las que se dispone de datos actuales (diarios). Cuando los datos externos (verdaderos básicos) tienen una naturaleza diaria y están actualmente disponibles para una muestra de aplicaciones, es posible estimar el número de descargas para una aplicación fuera de esta muestra con una clasificación  $r$  determinada utilizando cualquiera de los parámetros/ajustes modelos descritos anteriormente. El proceso para implementar este algoritmo puede tomar las siguientes etapas:

35

1. Para el modelo seleccionado, estimar el valor de todos los parámetros distintos de  $r$  suponiendo que los valores de  $r$  se toman de los datos externos (verdaderos básicos) y utilizando la versión logarítmica del modelo aplicable para estimar los parámetros restantes.

40

2. Repetir el modelo original seleccionado, aplicando los valores estimados de los parámetros de (1) en lugar de los parámetros indicados en el modelo.

3. Dada una aplicación con clasificación  $r$  en un día determinado, el número de descargas de esa aplicación en ese día determinado puede estimarse calculando la fórmula determinada en (2).

El algoritmo 2 se refiere a los escenarios en los que hay una disponibilidad discontinua de datos externos (verdaderos básicos), de manera que inicialmente hay datos diarios disponibles y, posteriormente, solo hay datos



mensuales disponibles. Entonces, todavía es posible estimar el número de descargas para una aplicación con una clasificación  $r$  determinada utilizando cualquiera de los modelos parametrizados/ajustados descritos anteriormente.

El procedimiento para implementar el algoritmo 2 incluye:

- 5 1. Estimar los parámetros de forma según corresponda calculando el promedio de los parámetros apropiados sobre un modelo diario preparado en base a los datos diarios para el período de tiempo en que dichos datos estaban disponibles.
2. Estimar el parámetro de escala  $a$ :
  - (a) Tomando los valores más altos de descargas como se ve en los datos externos de los dos meses anteriores,  $y$ ;
  - 10 (b) Dividiendo cada valor de (a) por 30 para determinar una estimación lineal del número de descargas diarias para cada día de los dos meses anteriores,  $y$ ;
  - (c) Usando los valores de (b), extrapolar linealmente el número de descargas  $d$  para la fecha actual (o la fecha en cuestión). Basado en  $d$  y si son necesarios los parámetros de forma de (1), estimar el parámetro  $a$ .
- 15 3. Repetir la fórmula del modelo original seleccionado, aplicando los valores de los parámetros de (1) y (2) en lugar de los parámetros en el modelo.
4. Dar una aplicación con clasificación  $r$  en un día determinado, el número de descargas de esa aplicación en ese día puede estimarse calculando la fórmula determinada en (3).

20 Con referencia a la figura 8 y el escenario representado en 800, la lógica 217 de aplicación del medidor de investigación que se ejecuta en los dispositivos del usuario está configurada para proporcionar una forma pasiva de observar/registrar el uso del panelista de los dispositivos digitales. Está específicamente destinado a observar, registrar y cargar 814 de forma pasiva un registro 812 de la actividad del panelista en el(los) servidor(es) del sistema para almacenamiento y análisis 816, 818.

La lógica 217 de medición reside en la memoria del dispositivo medido. Cuando se produce un evento 808 particular en el dispositivo, el medidor 217 se dispara 804 se informa sobre el evento 806, 810 registrado.

- 25 La administración y transferencia de archivos de registro pueden adherirse a una o más de las siguientes reglas y procedimientos:
- Cada archivo de registro se toma como huella digital a un panelista particular y un dispositivo medido específico,
  - Los archivos de registro se almacenan localmente en el dispositivo medido,
  - 30 • Los archivos de registro se comprimirán localmente en el dispositivo medido usando un algoritmo de compresión relevante,
  - Los archivos de registro se transfieren de forma asíncrona a un servidor remoto 'en espera' (802) en función de la disponibilidad de una conexión a Internet.
  - Para minimizar el impacto en el rendimiento del dispositivo, la transferencia de registros se iniciará solo cuando haya una conexión a Internet disponible, y/o
  - 35 • En el caso de que no haya conexión a Internet disponible durante un período de tiempo prolongado, se almacenarán localmente en el dispositivo medido, por ejemplo, aproximadamente 10 MB, de archivos de registro.

Los datos recogidos por el software 217 de medición puede ser dividida en múltiples, por ejemplo, tres, categorías conceptuales:

- 40
- Metadatos del evento: Esta es información que describe el contexto para el evento y es preferiblemente idéntica en cada tipo de evento registrado por el medidor de investigación;
  - Datos del evento: Esta es información específica del evento registrado y su contenido puede variar según el tipo de evento que se haya registrado; y/o
  - Datos del latido del corazón: Esta información informa al sistema que el panelista/dispositivo permanece "activo" y conectado a Internet.

45 Los datos de eventos registrados por el medidor 217 serán procesados preferiblemente de manera centralizada por el(los) servidor(es) del sistema para:

- 50
- a) limpiar los datos,
  - b) conectar los metadatos del evento, cuando sea necesario, a eventos subyacentes,
  - c) mapear identificadores a las entidades relevantes, y/o
  - d) calcular métricas a partir de los datos registrados.

Cada evento registrado por el medidor 217 de investigación preferiblemente registrará al menos algunos de los

siguientes metadatos, por ejemplo:

- Panelista/ID del dispositivo. Este es un identificador único que identifica a un panelista en particular y un dispositivo medido específico (panelista/huella digital del dispositivo).
- Marca de tiempo. Esta es una marca de tiempo UTC que indica cuándo ocurrió el evento.
- 5 • Geolocalización. Esta es información técnica sobre la ubicación geográfica del Usuario del Cliente cuando se produjo el evento. Si esto se puede capturar en el nivel de las coordenadas geográficas (mediante el GPS integrado o el software de ubicación inalámbrico), sería ideal. Este punto de datos puede aceptar un valor de "Desconocido".

10 El medidor 217 puede estar configurado para utilizar un servidor externo para recoger un número de puntos de datos seleccionados, preferiblemente aún siempre relacionado con el medidor en el dispositivo que se ejecuta en el propio dispositivo. Una configuración ejemplar de tal implementación será una VPN o un servidor proxy, a través del cual todos o algunos datos HTTP podrían dirigirse y registrarse por separado.

15 El medidor 217 en el dispositivo puede iniciar este tipo de configuración de perfil a través de la cual se puede usar el servidor externo para dichos fines de registro de datos, y el medidor en el dispositivo completará activamente al menos uno de los siguientes dos elementos:

1. Autenticar e identificar al usuario, utilizando una configuración de perfil personalizada o huellas dactilares ocultas en el tráfico, para facilitar la identificación del panelista o dispositivo, y hacer que sea comparable/relacional con los datos que el medidor del propio dispositivo está recogiendo directamente.
- 20 2. Marcas de tiempo coincidentes, mediante el uso de algunas trazas sincronizadas u ocultas, o la transferencia proactiva de paquetes HTTP, para generar un proceso a través del cual las marcas de tiempo de los datos de medición en el dispositivo y los datos recopilados en un servidor externo pueden coincidir perfectamente (esto implica la creación de los llamados valores de compensación entre el registro en el dispositivo y el servidor externo).

25 Una característica adicional del medidor 217 es un denominado controlador de datos de contexto, que está configurado para añadir información contextual a cualquier actividad de registro, independientemente de la actividad de explotación que tiene lugar en el propio contador de investigación o en un servidor independiente. Este controlador de contexto puede configurarse para agregar información de ubicación, información de rendimiento del dispositivo y/o información de estado de la pantalla en cualquier medida. Un propósito práctico de esta función es agregar información sobre la actividad del usuario y el estado de la pantalla a los datos recopilados desde servidores 30 externos (más ficticios), como servidores VPN o servidores proxy. Por lo tanto, el tráfico de uso activo se puede separar de las actividades de tráfico de fondo en cualquier dispositivo: uso activo significa que la pantalla está encendida y que el usuario está haciendo algo activo con el dispositivo.

35 La cantidad de información registrada para diferentes eventos depende de la naturaleza del evento que se está registrando. Los eventos recopilados pueden estar relacionados con el uso de aplicaciones, solicitudes de páginas de Internet, llamadas de voz realizadas, etc.

En su forma en bruto, los datos recogidos por el medidor 217 serán extremadamente granulares y pueden abarcar muchos miles de sitios web, aplicaciones móviles, sistemas operativos, agentes de usuario, y dispositivos móviles. La granularidad de estos datos crea dos problemas separados en el proceso de producción y entrega de datos:

- 40 • Alineación de granularidades de datos. Los datos que son esenciales para el proceso de producción de datos (es decir, datos del estudio estructural, datos de la tienda de aplicaciones, datos HTTP) no se recopilan al mismo nivel de granularidad que los datos recopilados por el medidor 217. Para calcular y aplicar los factores de escala/proporción derivados de estos componentes, los datos altamente granulares deben clasificarse para alinearlos con fuentes de datos menos granulares.
- 45 • Experiencia del cliente. Para que el entregable al cliente sea navegable y legible, es importante que los clientes cuenten con la capacidad de navegar y buscar las entidades específicas (es decir, sitios web, aplicaciones móviles, sistemas operativos, agentes de usuario, dispositivos móviles) que les interesen a ellos. Aunque una experiencia de búsqueda puede admitir datos altamente granulares, una experiencia de navegación de entregables, etc. a través de la interfaz de usuario del sistema no puede: sería imposible encontrar entidades específicas entre las decenas de miles de entidades contenidas a nivel granular.

50 Una forma factible de abordar con eficacia estos dos problemas es desarrollar un sistema para administrar los metadatos 214. Los metadatos que residirán dentro del sistema de administración de metadatos se utilizarán durante el proceso de producción de datos y serán esenciales para producir el conjunto final de datos entregables.

55 Los metadatos 214 mencionados anteriormente se utilizan para clasificar/organizar las entidades que son observables, por ejemplo, dentro de los archivos de registro del medidor de investigación, los datos 212 de la tienda de aplicaciones, los datos 211 del censo en general y/o los datos 208 del estudio estructural. Debido a que el marco descrito aquí abarca una variedad de entidades conceptuales, soportará simultáneamente una variedad de esquemas ontológicos. En general, estos esquemas se pueden dividir en dos categorías diferentes:

- Esquema univalente. Las entidades incluidas en un esquema univalente tienen un procedimiento de categorización canónica. Su jerarquía se define canónicamente, por ejemplo, por el personal del sistema que administra y actualiza el sistema y la(s) lógica(s) relacionada(s).
- Esquema multivalente. Las entidades asignadas a un esquema multivalente tendrán múltiples procedimientos de organización canónicos. Los clientes en diferentes audiencias probablemente tendrán diferentes enfoques para la categorización de la entidad dependiendo de sus necesidades comerciales particulares. Cada esquema multivalente será su propio procedimiento de organización y se considerará canónico según las reglas de edición definidas por los miembros del personal.

Debido a que muchas de las entidades medidas por el marco descrito aquí son de naturaleza multivalente, esto significa que habrá varios esquemas canónicos diferentes según los cuales el sistema descrito en esta invención publica sus datos.

Con referencia a la figura 9, cada uno de estos esquemas se puede derivar de una o más fuentes, como se deduce de la figura. Los editores 902 de taxonomía pueden incluir, por ejemplo, la entidad que implementa la metodología general. Uno o más esquemas desarrollados y mantenidos por el personal del sistema utilizando, por ejemplo, un sistema de gestión de metadatos a medida, podrían asumirse como el esquema canónico "predeterminado". Sin embargo, por ejemplo, las tiendas 914 de aplicaciones y los proveedores 920 de tecnología pueden proporcionar esquemas de categorización considerando, por ejemplo, el hecho de que las aplicaciones disponibles en las tiendas de aplicaciones generalmente han sido categorizadas por sus desarrolladores/la propia tienda de aplicaciones. Se puede establecer la taxonomía 906 del sistema. Aun así, los usuarios finales de las herramientas de informes proporcionadas, por ejemplo, pueden definir sus propias taxonomías y categorías 912, a través de la aplicación de la cual se les presentan los datos.

En general, en el contexto de la presente invención, la provisión de entregables fiables y objetivos sobre los temas deseados requiere aún la validación de los datos de panel y su calibración utilizando los medios disponibles.

Por lo tanto, preferiblemente de manera automática al recibir los datos del panel, el sistema está configurado para preprocesar los datos. En particular, se pueden seguir las siguientes etapas:

- Validación de los panelistas: los panelistas se validan utilizando las reglas de validación aplicables al panel 216, 218 en cuestión. Después de la validación, aquellos (es decir, los datos de) panelistas que han sido marcados como válidos pueden ser sometidos a más etapas de preprocesamiento.
- Cálculo de métricas y/o reorganización de datos. Para acelerar el cálculo de los entregables, los registros sin procesar de los panelistas validados se procesan para determinar métricas predeterminadas y/o para organizar los datos en una estructura propicia para un análisis posterior.
- Metadatos asociados. Paralelamente al cálculo de las métricas necesarias/reorganización de datos sin procesar, los registros se asocian con las entradas de metadatos correspondientes.

Después del preprocesamiento de los datos del panel, las acciones de calibración pueden ejecutarse utilizando datos como los datos 211 de nivel de censo y/o los datos de estratificación derivados de la muestra, opcionalmente estudio 208 estructural.

El proceso de calibración puede ser y típicamente es un proceso semimanual o semiautomático, que potencialmente requiere, por ejemplo, que un analista calcule varios valores, los ingrese en sistemas analíticos, ejecute procesos computacionales, analice los resultados y repita el proceso como sea necesario.

Conceptualmente, el proceso de calibración puede incluir fases de revisión de datos, cálculo del valor de control y aplicación de uno o más algoritmos de procesamiento, tal como un algoritmo de rastreo para determinar las ponderaciones de calibración adecuadas para los datos a nivel de usuario basados en los valores de control establecidos.

La figura 10 ilustra, solo a modo de ejemplo general, una realización de un procedimiento de revisión de muestras para la selección de la variable de calibración y de la categoría para permitir la determinación del valor de control. Un experto en la técnica entenderá que puede adaptar, por ejemplo, el procedimiento ilustrado, o utilizar uno alternativo, para cumplir mejor los objetivos de diseño establecidos por cada escenario de uso en la vida real de la presente invención.

De hecho, los datos 1002 de muestra disponibles pueden ser analizados y, por ejemplo, la siguiente indicación determinada:

- Frecuencias 1004 de categoría, la frecuencia de categorías individuales entre los panelistas validados (por ejemplo, género: varón como porcentaje del total de la muestra no ponderada) para las variables de calibración.
- Totales nominales/Medios 1006, el valor total de las variables de calibración nominales (es decir, las descargas totales de aplicaciones móviles) y el valor medio de las variables de calibración nominales.
- Recuento 1008 de casos de categoría, el recuento no ponderado de casos en cada categoría entre los panelistas validados (por ejemplo, el número de casos con género: varón) para todas las variables de calibración.

Según los cálculos anteriores, las categorías seleccionadas pueden colapsarse (frecuencia/recuento por debajo de un umbral predeterminado) para mejorar la estabilidad de la calibración, por ejemplo. Un conjunto optimizado de variables de calibración y categorías se formará 1010.

5 Finalmente, el sistema está configurado para calcular los valores 1012 de control en base a estas variables de calibración y categorías mediante la utilización de los datos 211 disponibles a nivel de censo y, por ejemplo, datos proporcionados por el estudio 208 estructural.

10 Tras la calibración, los entregables pueden someterse a una serie de controles de control de calidad seleccionados. Este proceso de control de calidad se puede realizar cuidadosamente, por ejemplo, mensualmente, después de la actualización de los datos de calibración del estudio estructural. Durante los procedimientos de control de calidad, tanto la estabilidad como la precisión de los datos podrían examinarse utilizando cualquier procedimiento adecuado según lo apreciado por un experto.

En última instancia, el(los) entregable(s) se puede(n) almacenar en al menos un repositorio de datos, tal como una base de datos en el sistema de fondo, y a los clientes se les puede conceder el acceso a elementos de datos específicos a través de la interfaz de usuario del sistema.

15 En particular, los datos pueden ser entregados a clientes suscritos a través de una interfaz basada en web. Esta interfaz basada en web puede cumplir con una o más de las siguientes características:

- Compatible con los principales navegadores (es decir, MSIE, Firefox, Chrome, Safari).
- Entrega sobre SSL.
- Funcionalidad/apariencia similar en todos los clientes (es decir, los datos disponibles/contenido pueden diferir entre los grupos de clientes, pero la herramienta sigue siendo la misma).

20 Anteriormente se han descrito varias realizaciones de la presente invención en detalle. Estas realizaciones se relacionaron generalmente con dispositivos móviles e Internet y otros dispositivos similares conectados a redes de comunicaciones inalámbricas o por cable, y mostraron cómo los dispositivos móviles e Internet comúnmente disponibles pueden aprovecharse para recopilar datos de forma automática y pasiva en el contexto del terminal, como parámetros o rendimiento de la red o dispositivo, o sobre el comportamiento de un usuario en relación con el uso de ese dispositivo en particular u otros dispositivos o redes circundantes, y en última instancia, cómo todos estos datos se pueden canalizar a un sistema que crea en caché o en tiempo real los análisis y los diferentes tipos de información derivada basándose en los datos recopilados, y se pueden combinar esos datos con datos de otros dispositivos o datos relacionados con otras personas, y finalmente presentar los datos cultivados a través de varias aplicaciones de vuelta al usuario u otros destinatarios.

30 Sin embargo, una persona experta apreciará el hecho de que el alcance no se limita a las realizaciones divulgadas de manera explícita, sino que abarca también otras realizaciones que caen bajo la redacción de las reivindicaciones.

35 En diversas realizaciones del sistema, el sistema puede incluir, además de un número de servidores, una pluralidad de dispositivos de usuario y, además, opcionalmente, un número de datos a nivel de censo que proporcionan fuentes de datos centradas en la red.

En diversas realizaciones, el esquema de metadatos ontológicos mencionado incorpora opcionalmente una base de datos de metadatos unificados, que describe interrelaciones y proporciona para hacer coincidir los conceptos y los sujetos, aunque mide y se refleja de manera diferente en diferentes plataformas y fuentes de datos que proporcionan los datos recogidos, pero refleja los mismos temas.

40 En varias realizaciones, el esquema de metadatos incorpora una pluralidad de esquemas canónicos para categorizar y organizar datos, en el que opcionalmente hay un número de esquemas univalentes y multivalentes, además opcionalmente uno o más de los esquemas canónicos para producir un entregable que se selecciona dinámicamente por el destinatario entregable.

45 En varias realizaciones, el dispositivo maestro mencionado anteriormente está configurado opcionalmente para ejecutar al menos una acción seleccionada del grupo que consiste en: seguimiento y registro de tráfico de datos, señales u otros eventos que tienen lugar entre el(los) dispositivo(s) conectado(s) y el dispositivo maestro, encuestar de forma pasiva o activa las interfaces de programación de aplicaciones (API) de uno o más dispositivos conectados funcionalmente para recopilar datos de medición, almacenar datos de medición recopilados temporalmente o durante un período de tiempo predeterminado, y enviar datos a una o más redes en base a servidores de recogida de datos para su análisis.

50 En diversas realizaciones, los datos del nivel de censo mencionados anteriormente pueden obtenerse de una fuente de datos centrada en la red de Internet, que indica la distribución y/o la escala de las características del sujeto y que incluye opcionalmente puntos de datos indicativos del comportamiento de una o más poblaciones definidas de usuarios, dispositivos, aplicaciones, servicios o sitios web durante un período de tiempo predeterminado, opcionalmente a diario. Opcionalmente, los datos a nivel censales mencionados anteriormente se obtienen desde o a través de varias fuentes de datos centradas en la red de Internet que comprenden al menos un elemento

seleccionado del grupo que consiste en: operador de red, red de anuncios, intercambio de anuncios, plataforma de análisis de sitios o aplicaciones, pasarela de red, proxy de red, plataforma de administración de contenido, tienda de aplicaciones (APP) y otra plataforma de distribución o entrega de contenido.

- 5 Opcionalmente, los datos a nivel de censo mencionados con respecto a varias aplicaciones se obtienen desde al menos una plataforma de distribución de contenido que incluye opcionalmente una tienda de aplicaciones o un mercado digital.

**REIVINDICACIONES**

1. Sistema electrónico (100, 108, 110, 400, 401A, 401B) para obtener y analizar datos sobre el comportamiento de un usuario (102) en línea y otra actividad relacionada con dispositivos de usuario que pueden conectarse a Internet, opcionalmente dispositivos móviles, comprendiendo el sistema una serie de dispositivos de servidor (108) conectados funcionalmente que comprenden un dispositivo de procesamiento (420) para procesar instrucciones y datos, una memoria (422) para almacenar instrucciones y otros datos y una interfaz de datos (426) para transferir datos, estando configurado dicho sistema para
- 5                   recopilar (404, 408) datos desde una pluralidad de fuentes de datos, siendo dichos datos recopilados indicativos de Internet (107), contenido, medios, aplicaciones y/o uso del dispositivo, que comprenden
- 10                  datos individuales a nivel de usuario adquiridos a partir de dispositivos de usuario (104, 104a, 104b, 105, 217, 402) a través de medición pasiva, recopilándose dichos datos a nivel de usuario preferiblemente usando al menos un panel de investigación de usuarios (216) de constitución controlada, y
- datos a nivel de censo (107a, 211) que indican el comportamiento y las características demográficas en toda la población de usuarios, dispositivos activos o servicios medidos,
- 15                  organizar (406) los datos a nivel de usuario obtenidos y los datos a nivel de censo en un conjunto de datos categorizados, preferiblemente multivalentes, usando un esquema de metadatos ontológicos (214, 414), y
- determinar un producto electrónico (112, 206, 416) basado en una integración de los datos a nivel de usuario y los datos a nivel de censo, en donde los datos a nivel de censo se usan para calibrar los datos a nivel de usuario,
- 20                  teniendo el producto un alcance seleccionable dinámicamente en términos de una serie de factores de interés con respecto a dispositivos usados o comportamientos en línea, incluyendo preferiblemente el uso de la aplicación, la distribución de la aplicación, el uso del contenido, la distribución del contenido, la monetización de la aplicación, la demografía del usuario, la distribución del dispositivo, las características del dispositivo, el uso del dispositivo y/o
- 25                  factores de tiempo, **caracterizado porque** el sistema está configurado para recopilar dichos datos a nivel de usuario a través de uno o más dispositivos maestros (104, 300) de un usuario que tiene una pluralidad de dispositivos de usuario, en donde hay una serie de otros dispositivos de usuario (104a, 104b) conectados funcionalmente a dicho
- 30                  uno o más de dichos dispositivos maestros, midiéndose el uso de dichos otros dispositivos de usuario mediante los dispositivos maestros correspondientes, actuando los dispositivos maestros como nodos de recopilación de datos y conexión hacia el sistema,
- y, además, en donde el sistema está configurado para distinguir, en el análisis que da como resultado el producto, entre datos a nivel de usuario que se originan en diferentes dispositivos de usuario de la pluralidad de dispositivos de usuario del usuario basado en los identificadores de dispositivo asociados a los datos a nivel de usuario
- adquiridos a partir de los dispositivos del usuario.
2. El sistema de la reivindicación 1, en el que los datos recopilados incluyen además datos de estratificación estadística derivados de la muestra (208) que estiman la distribución de las características demográficas en la
- 35                  población de usuarios, estando el sistema configurado para organizar también los datos de estratificación en dicho conjunto de datos y usarlos en la determinación del producto, incluyendo la calibración de datos a nivel de usuario.
3. El sistema de cualquier reivindicación anterior, configurado para recopilar (217) datos a nivel de usuario desde un primer panel de investigación de usuarios (216) de constitución controlada y que tiene un primer número de participantes, y un segundo panel de investigación de usuarios (218) con requisitos de inscripción reducidos y que
- 40                  tiene un mayor número de participantes, en donde el sistema está configurado opcionalmente para calibrar datos a nivel de usuario obtenidos a través del segundo panel de investigación de usuarios usando datos a nivel de censo, opcionalmente datos de estratificación derivados de la muestra y datos a nivel de usuario obtenidos por medio del primer panel de investigación de usuarios.
4. El sistema de cualquier reivindicación anterior, configurado para usar el esquema de metadatos para alinear datos
- 45                  a nivel de usuario más granulares con datos a nivel de censo menos granulares a través de categorización de datos.
5. El sistema de cualquier reivindicación anterior, configurado para recopilar datos a nivel de usuario usando al menos un elemento seleccionado del grupo que consiste en: aplicación de medición y registro (217) instalada en los dispositivos de usuario, opcionalmente dispositivos terminales inalámbricos, y lógica de medición y registro (217) integrada con software destinado a otro uso, incluyéndose dicha lógica con el software.
- 50                  6. El sistema de cualquier reivindicación anterior, configurado para recopilar datos adicionales a nivel de usuario indicativos de eventos de conducta o contextuales que tienen lugar de manera externa a un dispositivo de usuario obtenidos a través de un sensor, opcionalmente una cámara o un micrófono, del dispositivo.
7. El sistema de cualquier reivindicación anterior, en el que la lógica de medición y registro en un dispositivo de usuario (104, 104a, 104b, 217, 402) para permitir que la recopilación de datos a nivel de usuario comprenda al
- 55                  menos una característica seleccionada del grupo que consiste en: recopilador de datos de latido indicativo de los periodos de actividad del usuario o del dispositivo y/o periodos de conectividad relacionados con Internet, autenticación de usuario, identificación de usuario, registro de identidad de usuario, registro de datos de eventos, metadatos de eventos o registro de datos de contexto, registro de marca de tiempo, registro de geolocalización,

registro de estado de WiFi, registro de estado en línea, registro de operador, registro de estado de itinerancia, registro de modo de orientación, registro de datos de descarga, registro de datos de carga, comparación y registro de datos en el dispositivo y datos proporcionados por dispositivos externos.

- 5 8. El sistema de cualquier reivindicación anterior, configurado para conectarse a al menos una tienda (602, 604) de aplicaciones u otra plataforma de distribución de contenidos remota que usa un rastreador o una entidad lógica de robot configurada, para obtener dichos datos a nivel de censo, para ejecutar al menos una acción seleccionada del grupo que consiste en: emular un usuario o un dispositivo de usuario asociado a la plataforma, aplicar una API (interfaz de programación de aplicaciones) relacionada con la plataforma, conectarse a un proxy (606) entre el sistema y la plataforma, y aplicar una técnica de tunelado VPN.
- 10 9. El sistema de cualquier reivindicación anterior, configurado para obtener datos a nivel de censo con respecto a varias aplicaciones de al menos una plataforma de distribución de contenidos a través de la estimación de (700) el número de descargas de una aplicación basado en la posición de clasificación de las mismas, opcionalmente el número diario de descargas que se estiman basándose en la información de posición de clasificación preferiblemente diaria, y un modelo que aplica la posición de clasificación para proporcionar la estimación, en donde el modelo usado incluye preferiblemente al menos un modelo seleccionado del grupo que consiste en: modelo exponencial, modelo log-normal, modelo basado en ley de potencia con corte exponencial, y modelo Lomax.
- 15 10. El sistema de cualquier reivindicación anterior, configurado para obtener datos a nivel de censo con respecto a varias aplicaciones de al menos una plataforma de distribución de contenidos y para usar dichos datos a nivel de censo obtenidos desde dicha al menos una plataforma de distribución de contenidos para el control de calidad de estimaciones de comportamiento derivadas basándose en datos a nivel de usuario y/o para calibrar los datos a nivel de usuario.
- 20 11. El sistema de cualquier reivindicación anterior, configurado para determinar una serie de factores de calibración demográficos y de comportamiento basados en los datos a nivel de censo para los datos a nivel de usuario para la calibración.
- 25 12. El sistema de cualquier reivindicación anterior, configurado para estimar la base de usuarios de dispositivos de usuario y, opcionalmente, cambiar su tasa en función de los datos del censo, incluyendo preferiblemente datos centrados en la red de Internet.
- 30 13. Un procedimiento (1100) para obtener y analizar datos sobre el comportamiento de usuarios en línea y otra actividad relacionada con dispositivos de usuario que pueden conectarse a Internet, opcionalmente dispositivos móviles, que debe realizarse mediante un sistema electrónico que comprende al menos un servidor, comprendiendo dicho procedimiento
- recopilar (1106, 1108) datos desde una pluralidad de fuentes de datos, siendo dichos datos recopilados indicativos de Internet, contenido, medios, aplicaciones y/o uso de los dispositivos, incluyendo al menos
- 35 datos individuales a nivel de usuario adquiridos desde dispositivos de usuario a través de medición pasiva, recopilándose dichos datos a nivel de usuario preferiblemente usando al menos un panel de usuario de constitución controlada, y datos a nivel de censo que indican características demográficas y de comportamiento a través de toda la población de usuarios, dispositivos activos o servicios medidos,
- 40 organizar (1110) los datos obtenidos a nivel de usuario, los datos a nivel de censo y, opcionalmente, datos de estratificación derivados de la muestra en un conjunto de datos categorizados multivalente que usan un esquema de metadatos ontológicos, y determinar un producto electrónico basado en una integración de los datos a nivel de usuario, datos a nivel de censo y, opcionalmente, datos de estratificación derivados de la muestra, en donde los datos a nivel de censo y opcionalmente dichos datos de estratificación derivados de la muestra se usan para calibrar los datos a nivel de usuario, en donde el producto (1112) tiene un alcance seleccionable dinámicamente en términos de una serie de factores de interés con respecto a dispositivos usados o comportamientos en línea, incluyendo preferiblemente el uso de aplicaciones, la distribución de aplicaciones, el uso de contenidos, la distribución de contenidos, la monetización de aplicaciones, los datos demográficos del usuario, la distribución de dispositivos, las características del dispositivo, el uso del dispositivo y/o factores de tiempo, **caracterizado porque**
- 45 dichos datos a nivel de usuario se recopilan (300) a través de uno o más dispositivos maestros (104) de un usuario que tiene una pluralidad de dispositivos de usuario, en donde hay una serie de otros dispositivos de usuario (104a, 104b) conectados funcionalmente a dicho uno o más de dichos dispositivos maestros, el uso de dichos otros dispositivos de usuario se mide mediante los dispositivos maestros correspondientes, actuando los dispositivos maestros como nodos de recopilación de datos y conexión hacia el sistema,
- 50 y, además, en el que se distingue, mediante el sistema en el análisis que da como resultado el producto, entre datos a nivel de usuario que se originan en diferentes dispositivos de usuario de la pluralidad de dispositivos de usuario del usuario basado en los identificadores de dispositivo asociados a los datos a nivel de usuario adquiridos a partir de los dispositivos del usuario.
- 55

14. Un programa informático que comprende unos medios de código adaptados, cuando se ejecutan en un ordenador, para ejecutar el procedimiento de la reivindicación 13.

15. Un medio portador que comprende el programa informático de la reivindicación 14.



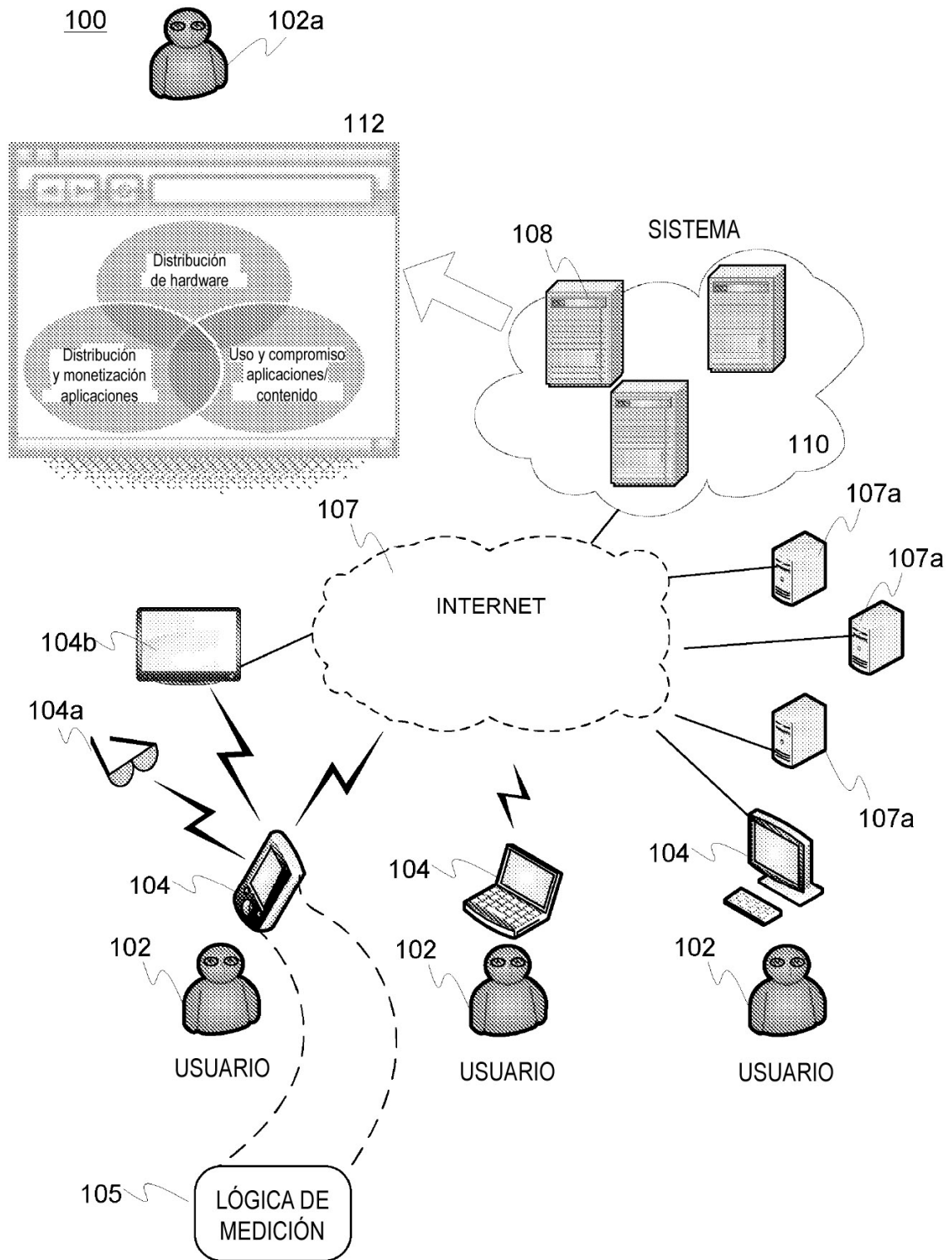


Figura 1

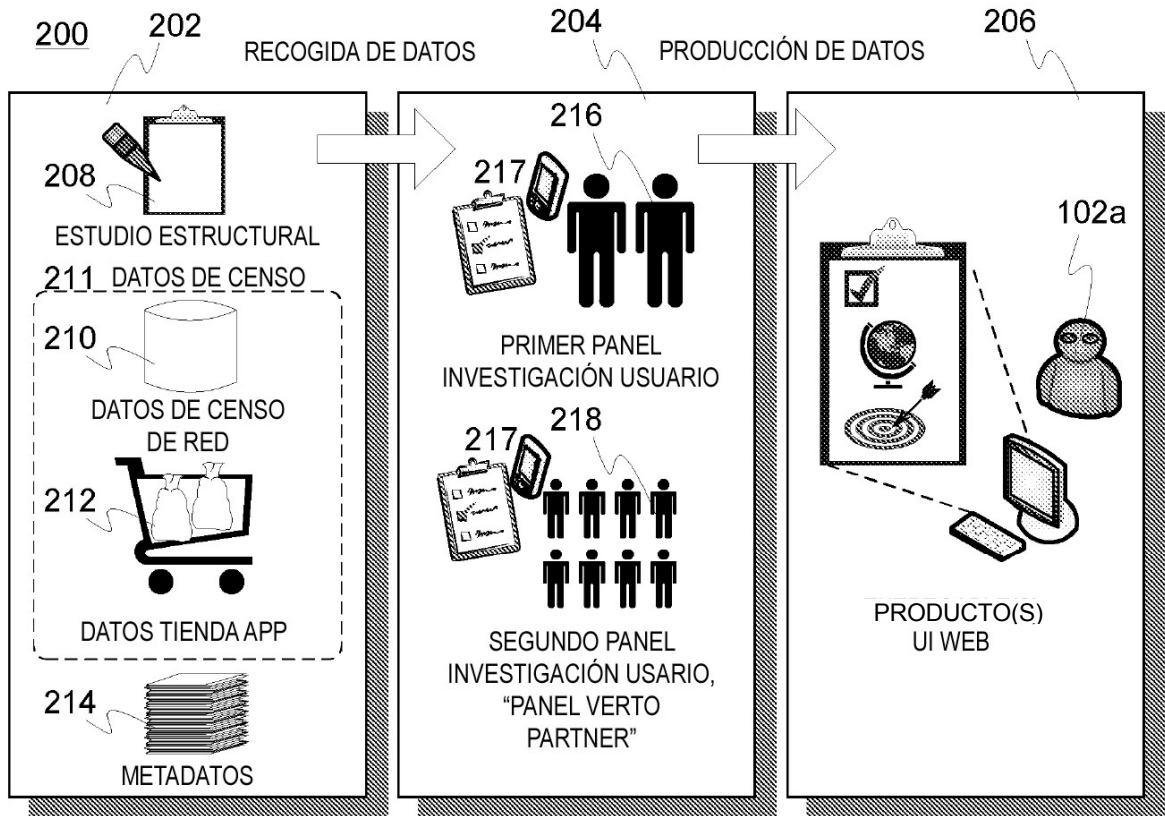


Figura 2

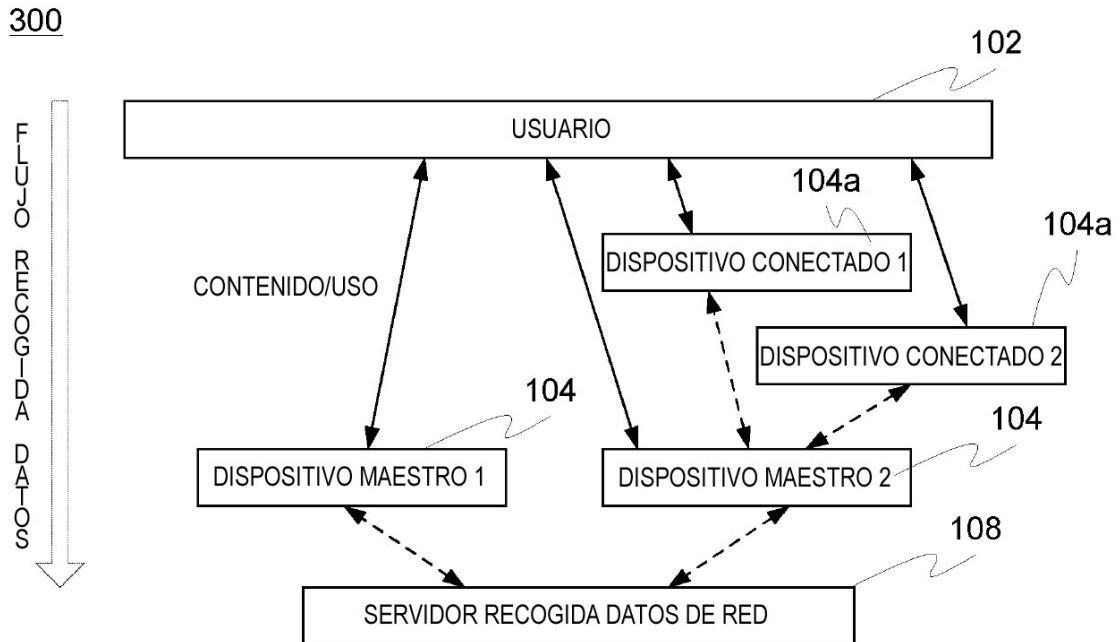


Figura 3

400

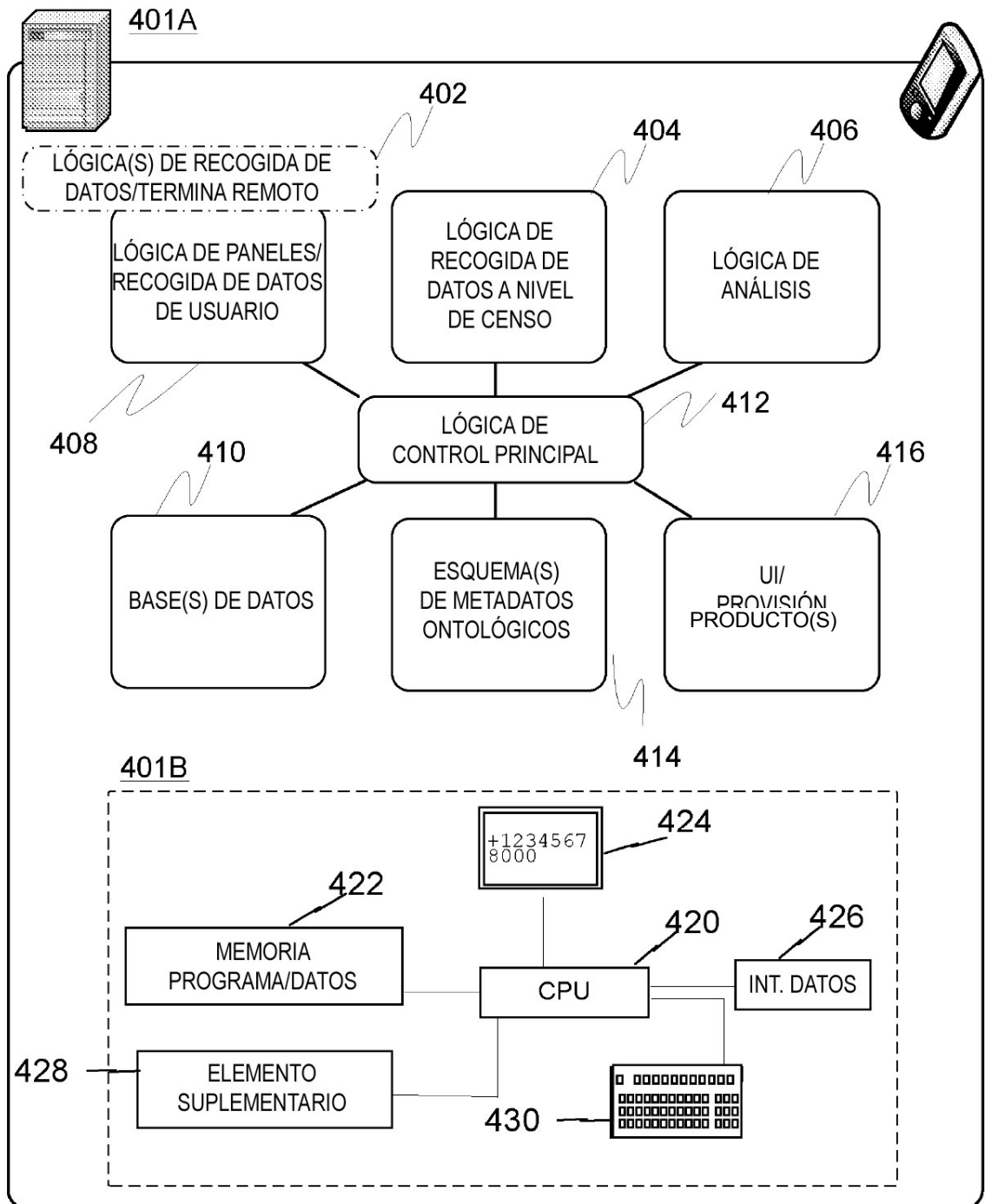


Figura 4

500

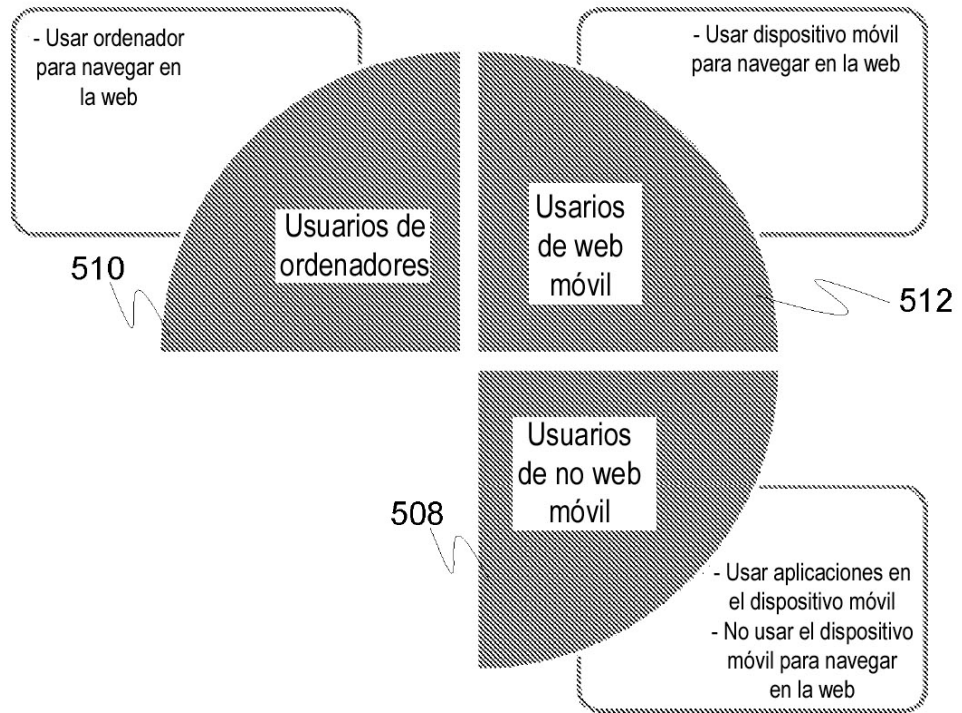


Figura 5

600

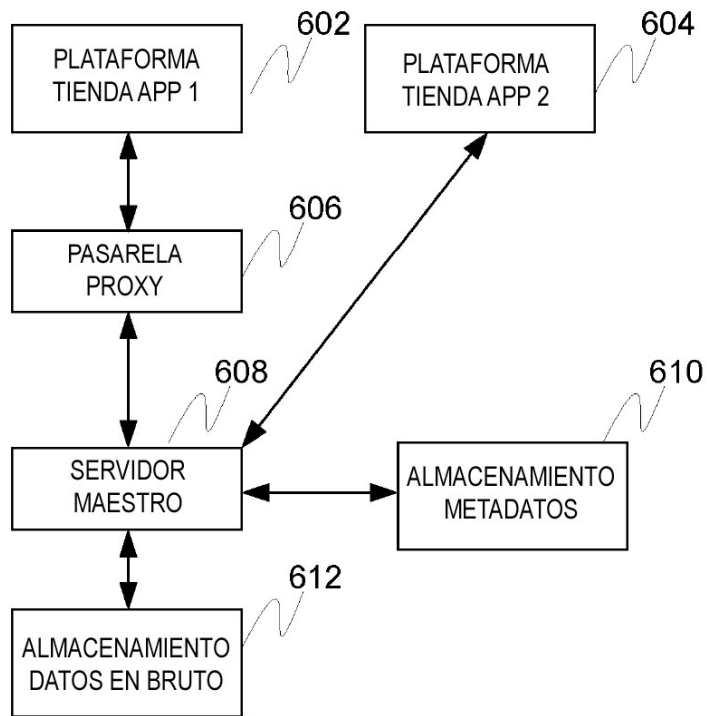


Figura 6

700

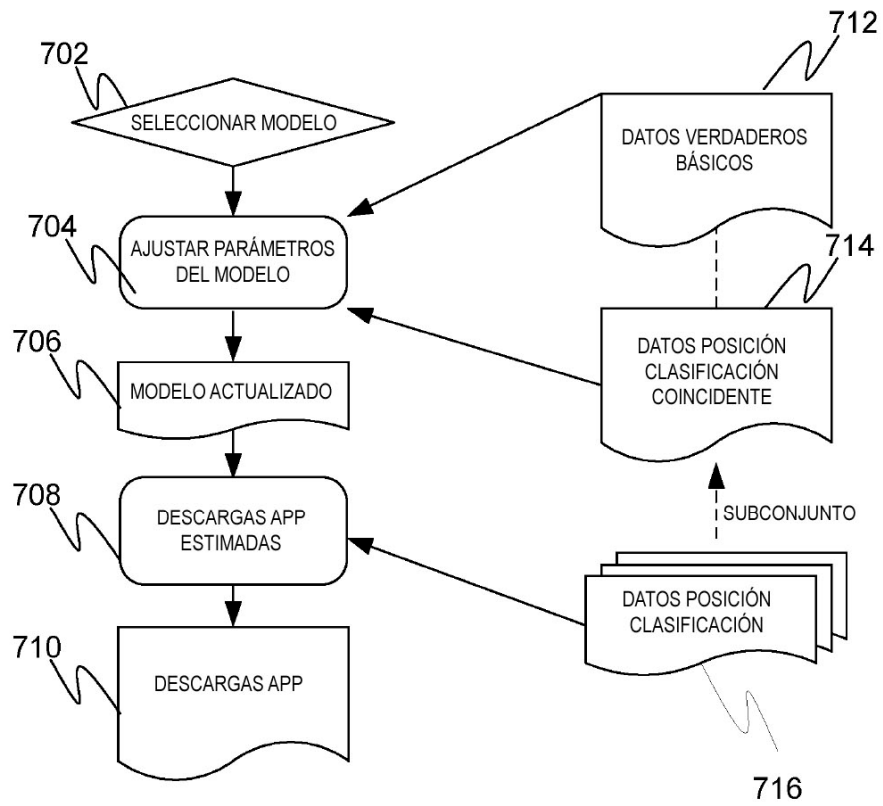


Figura 7

800

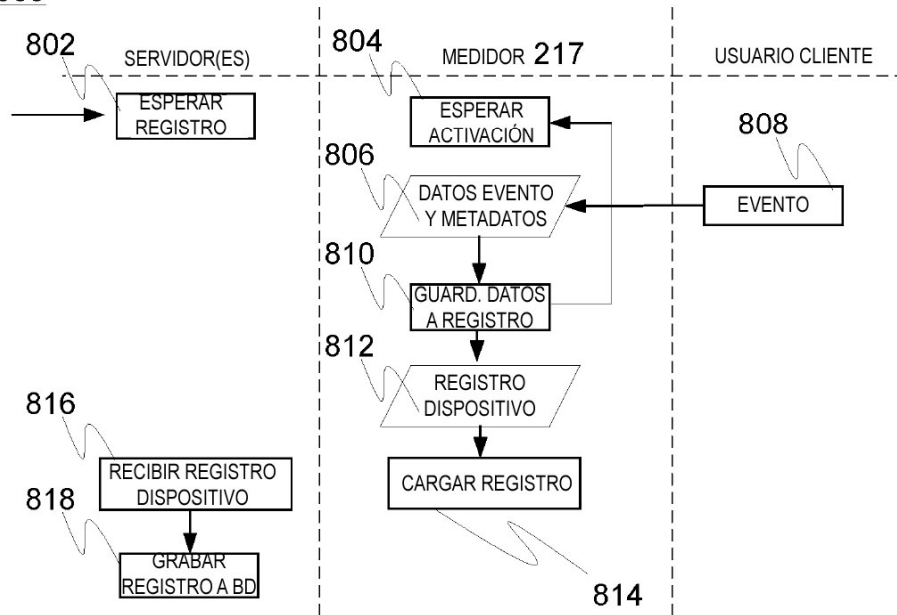


Figura 8

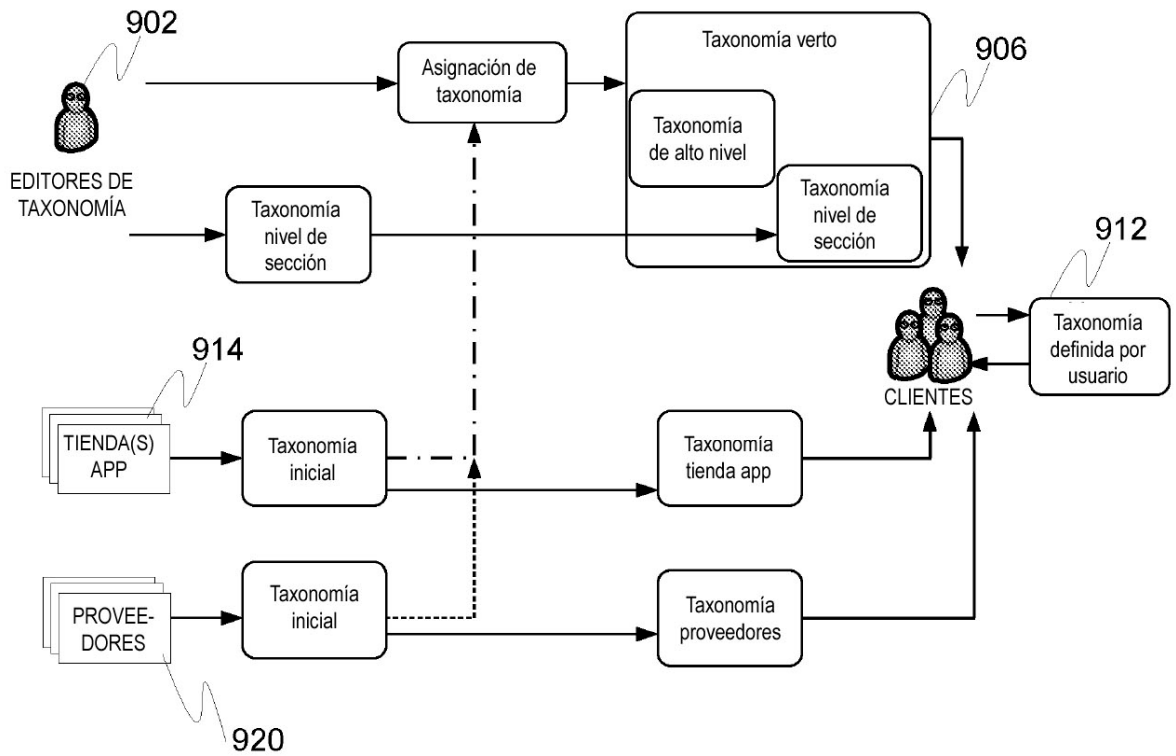


Figura 9

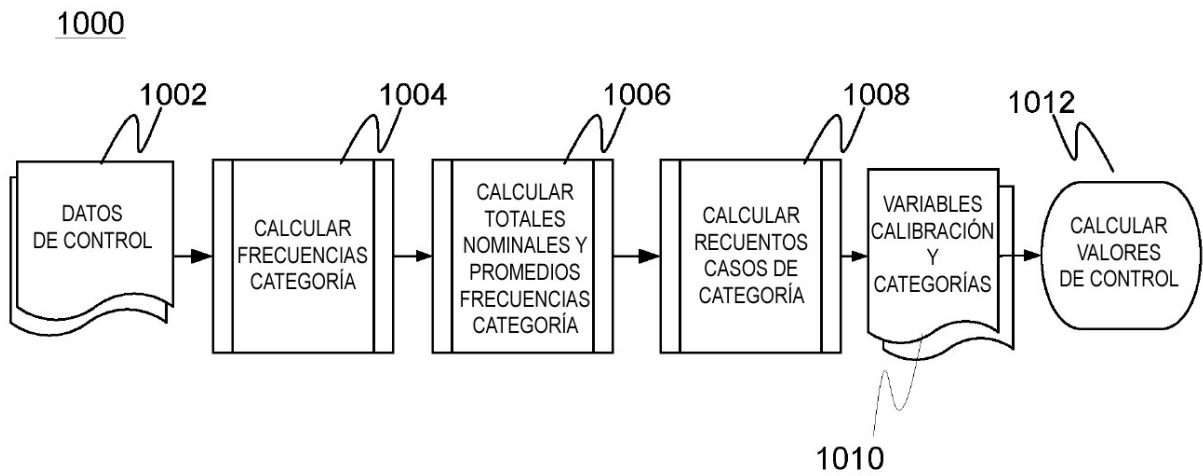


Figura 10

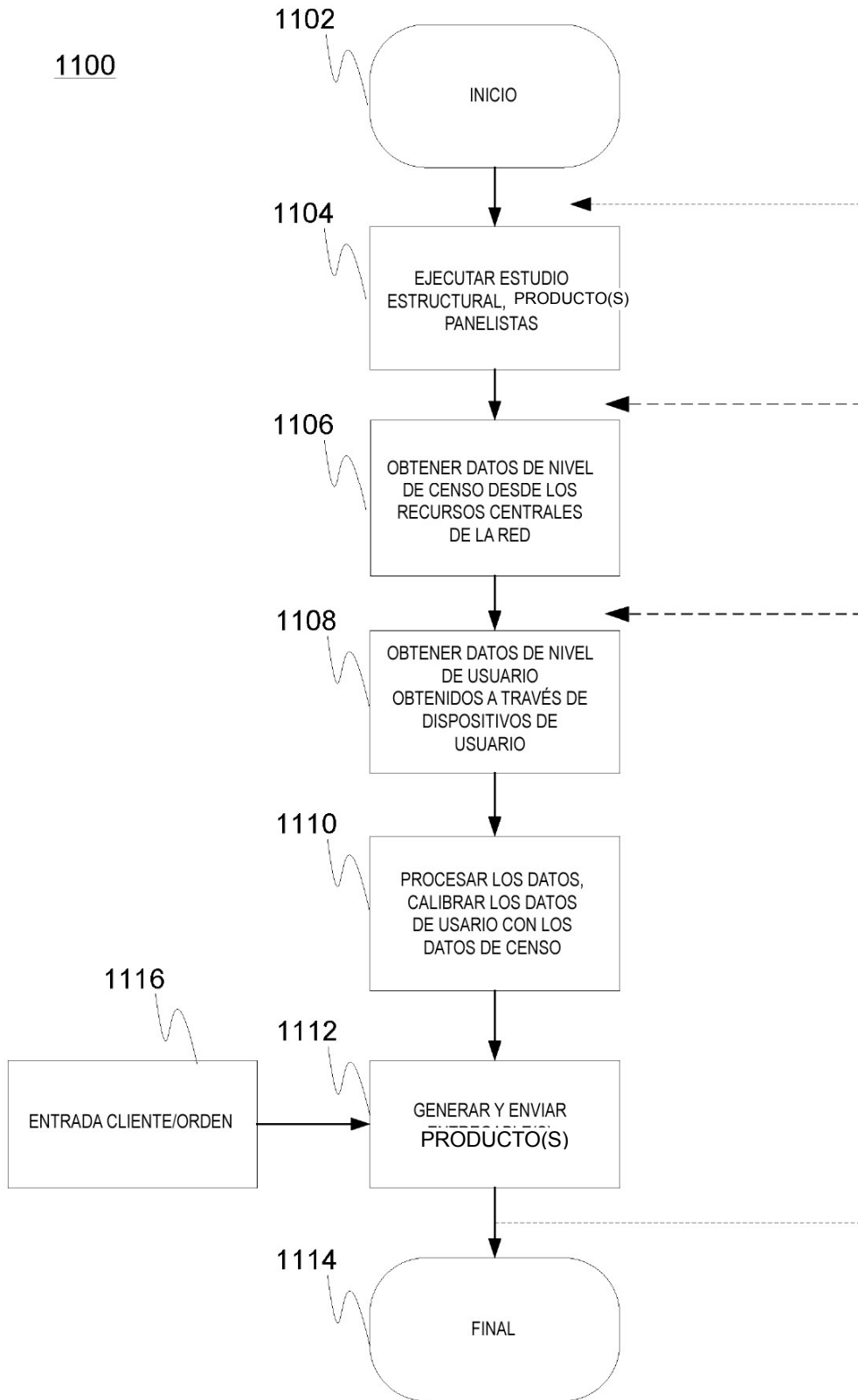


Figura 11