

**ACTUALIZACIÓN DE MODELOS DE CLASIFICADOR DE
ENTENDIMIENTO DE LENGUAJE PARA UN ASISTENTE DIGITAL
PERSONAL BASÁNDOSE EN EXTERNALIZACIÓN MASIVA**

5

ANTECEDENTES

A medida que la tecnología de cómputo ha avanzado, cada vez más potentes dispositivos móviles se han hecho en disponibles. Por ejemplo, los teléfonos inteligentes y otros dispositivos de cómputo se han hecho habituales. Las capacidades de procesamiento de esos dispositivos han dado lugar a diferentes tipos de funcionalidades que se están desarrollando, tales como funcionalidades relacionadas con asistentes digitales personales.

Un asistente digital personal puede utilizarse para realizar tareas o servicios para un individuo. Por ejemplo, el asistente digital personal puede ser un módulo de *software* se ejecuta en un dispositivo móvil o una computadora de escritorio. Además, un asistente digital personal implementado en un dispositivo tiene entendimiento conversacional interactivo e incorporado que es capaz de responder a las preguntas del usuario o los comandos de voz. Ejemplos de tareas y servicios que pueden ser realizados por el asistente digital personal pueden incluir la realización de llamadas telefónicas, enviar un correo electrónico o un mensaje de texto y configurar recordatorios de calendario.

Mientras un asistente digital personal puede ser implementado

para realizar múltiples tareas utilizando agentes, que programan/definen cada agente reactivo puede llevarle mucho tiempo. Por lo tanto, existe una amplia oportunidad para la mejora en las tecnologías relacionadas con la creación y edición de modelos de clasificador de entendimiento de definiciones de agente reactivo y lenguaje asociado para la implementación de un asistente digital personal.

BREVE DESCRIPCIÓN DE LA INVENCION

10

Esta Breve Descripción se proporciona para introducir una selección de conceptos en una forma simplificada que se describen a continuación en la Descripción Detallada. Esta Breve Descripción no pretende identificar características clave o características esenciales de la materia objeto reclamada, ni está destinada a ser utilizada para limitar el alcance de la materia objeto reclamada.

De acuerdo con uno o más aspectos, un método para actualizar modelos de clasificador de entendimiento de lenguaje puede incluir recibir, a través de uno o más micrófonos de un dispositivo de cómputo, una entrada digital de voz de un usuario del dispositivo de cómputo. La entrada también puede ser recibida de un usuario utilizando a través de otras entradas también, (por ejemplo, mediante entrada de texto u otros tipos de entrada). El procesamiento de lenguaje natural se realiza utilizando la entrada de voz digital para determinar una solicitud de voz de usuario. Al determinar la solicitud

de voz de usuario no coincide con al menos uno de una pluralidad de tareas predefinidas en la definición de un agente (por ejemplo, una definición de esquema de lenguaje de marcación extensible (XML) de un asistente digital personal que se ejecutan en el dispositivo de

5 cómputo, una interfaz gráfica de usuario de una herramienta de etiquetado de usuario final (EULT, según sus siglas en inglés) del dispositivo de cómputo puede ser utilizada para recibir una selección de un usuario. Una tarea puede ser definida por un comando de voz (o introducida por texto), así como por uno o más medios

10 adicionales, tales como a través de un motor basado en reglas, clasificadores de aprendizaje por máquina y así sucesivamente. La selección de usuario puede incluir al menos un intento de una pluralidad de intentos disponibles para un dominio. Opcionalmente, la selección de usuario también puede incluir de al menos una ranura

15 para al menos un intento. El al menos un intento está asociado con al menos una acción utilizada para realizar al menos una función de una categoría de funciones para el dominio. Cuando se incluye en la selección de usuario, la al menos una ranura indica un valor utilizado para realizar al menos una acción. Un conjunto de datos etiquetados

20 puede ser generado al emparejar (o de otra manera asociar) la solicitud de voz de usuario con la selección de usuario (por ejemplo, dominio seleccionado, intento y/o ranura). Un clasificador de entendimiento de lenguaje puede ser seleccionado de una pluralidad de clasificadores de entendimiento de lenguaje disponibles

25 asociados con la definición de agente, la selección basándose por lo

menos en al menos un intento seleccionado por el usuario. El clasificador de entendimiento de lenguaje seleccionado puede ser actualizado basándose el conjunto de datos etiquetados generados.

De acuerdo con uno o más aspectos, una computadora de
5 servidor que incluye una unidad de procesamiento y memoria
acoplada a la unidad de procesamiento. La computadora de servidor
puede configurarse para realizar las operaciones para actualizar
modelos de clasificador de entendimiento de lenguaje. Las
operaciones pueden incluir recibir de al menos un dispositivo de
10 cómputo de una pluralidad de dispositivos de cómputo
comunicativamente acoplados a la computadora de servidor, una
primera selección de usuario de al menos un intento de una
pluralidad de intentos disponibles. Opcionalmente, la selección de
usuario también puede incluir al menos una ranura para el al menos
15 un intento. Cuando se incluye en la selección de usuario, el al menos
un intento puede estar asociado con al menos una acción utilizada
para realizar al menos una función de una categoría de funciones
para un dominio. La al menos una ranura puede indicar un valor que
se utiliza para realizar al menos una acción. La primera selección de
20 usuario puede estar asociada con una entrada de voz digital recibida
en al menos un dispositivo de cómputo. Una pluralidad de
selecciones de usuario subsiguientes que son idénticas a la primera
selección de usuario puede ser recibida de al menos otro dispositivo
de cómputo de la pluralidad de dispositivos de cómputo. Un conjunto
25 de datos etiquetados puede ser generado al emparejar la entrada de

voz digital con la primera selección de usuario. Un clasificador de entendimiento de lenguaje puede ser seleccionado de una pluralidad de clasificadores de entendimiento de lenguaje disponibles asociados con una o más definiciones de esquema XML, la selección basándose al menos en una o más de la entrada digital de voz, el dominio, el intento y/o en la ranura de la primera selección de usuario. El clasificador de entendimiento de lenguaje seleccionado puede ser actualizado según el denominado conjunto de datos etiquetados generado.

De acuerdo con uno o más aspectos, un medio de almacenamiento legible por computadora puede incluir instrucciones que tras su ejecución hacen que un dispositivo de cómputo realice operaciones de actualización de los modelos de clasificador de entendimiento de lenguaje. Las operaciones pueden incluir determinar una solicitud de usuario basándose en la entrada de usuario recibida en el dispositivo de cómputo. La solicitud de usuario puede ser recibida a través de al menos uno de entrada de texto y entrada de voz y la solicitud puede ser para una funcionalidad de un asistente digital personal que se ejecuta en el dispositivo de cómputo. Las operaciones pueden incluir además determinar que la solicitud de usuario no coincide con al menos una de una pluralidad de tareas predefinidas (por ejemplo, comandos de voz) en una definición de esquema de lenguaje de marcación extensible (XML) del asistente digital personal. En una implementación, una puntuación de confianza puede ser generada mediante la aplicación

de una pluralidad de clasificadores de entendimiento de lenguaje disponibles asociados con la definición de esquema de XML para la solicitud de usuario. Tras determinar que la puntuación de confianza es menor que un valor de umbral, la selección de un usuario puede ser recibida mediante una interfaz gráfica de usuario de una herramienta de etiquetado de usuario final (EULT) del dispositivo de cómputo. En otra implementación, se pueden utilizar otros métodos (por ejemplo, en lugar de utilizar un valor de umbral) para determinar si se va a utilizar la EULT para recibir una selección de usuario de al menos uno de un dominio, un intento y/o información de ranura. La selección de usuario puede incluir al menos un intento de una pluralidad de intentos disponibles. Opcionalmente, la selección de usuario puede incluir un dominio y/o al menos una ranura para al menos un intento. El al menos un intento está asociado con al menos una acción utilizada para realizar al menos una función de una categoría de funciones para un dominio. Cuando se incluye en la selección de usuario, la al menos una ranura puede indicar un valor que se utiliza para realizar al menos una acción. Un conjunto de datos etiquetados puede ser generado al emparejar la solicitud de voz de usuario y la selección de usuario. Un clasificador de entendimiento de lenguaje puede ser seleccionado de la pluralidad de clasificadores de entendimiento de lenguaje disponibles asociados con la definición de esquema de XML, la selección basándose en al menos un intento y/o una ranura seleccionados por el usuario. Un clasificador de entendimiento de lenguaje actualizado

puede ser generado al entrenar al clasificador de entendimiento de lenguaje seleccionado mediante el denominado conjunto de datos etiquetados generado (por ejemplo, asociando el clasificador con la solicitud de voz y al menos uno del dominio, intento y/o ranura en la
5 selección de usuario).

Tal como se describe aquí, una variedad de otras características y ventajas pueden incorporarse en las tecnologías como desee.

10 **BREVE DESCRIPCIÓN DE LOS DIBUJOS**

La Figura 1 es un diagrama de bloques que ilustra un ejemplo de una arquitectura para actualizar modelos de clasificador de entendimiento de lenguaje, de acuerdo con una modalidad ilustrativa
15 de la descripción.

La Figura 2 es un diagrama de bloques que ilustran los varios usos de clasificadores de entendimiento de lenguaje por aplicaciones activados por voz, de acuerdo con una modalidad ilustrativa de la descripción.

20 Las Figuras 3A-3B muestran ciclos de procesamiento ilustrativos para actualizar modelos de clasificador de entendimiento de lenguaje, de acuerdo con una modalidad ilustrativa de la descripción.

Las Figuras 4A-4B ilustran interfaces de usuario ilustrativas de
25 una herramienta de etiquetado para usuario final, que se puede

utilizar de acuerdo con una modalidad ilustrativa de la descripción.

Las Figuras 5-7 son diagramas de flujo que ilustran la actualización de modelos de clasificador de entendimiento de lenguaje, de acuerdo con una o varias modalidades.

5 La Figura 8 es un diagrama de bloques que ilustran un dispositivo de cómputo móvil junto con innovaciones descritas aquí que pueden ser implementadas.

La Figura 9 es un diagrama de un sistema de cómputo ilustrativo, en el que algunas modalidades pueden ser
10 implementadas.

La Figura 10 es un entorno de cómputo de nube ilustrativo que puede utilizarse en conjunción con las tecnologías aquí descritas.

DESCRIPCIÓN DETALLADA

15

Como se describe aquí, diversas técnicas y soluciones pueden aplicarse para actualizar modelos de clasificador de entendimiento de lenguaje. Más específicamente, una especificación de definición de agente (por ejemplo, una especificación de definición de comando de voz (VCD), una especificación de definición de agente reactivo (RAD), u otro tipo de un documento legible por computadora) puede
20 utilizarse para definir uno o más agentes asociados con un asistente digital personal que se ejecuta en un dispositivo de cómputo. La especificación de definición de agente puede especificar la
25 información de dominio, información de intento, información de

ranura, información de estado, expresiones de usuario esperadas (o comandos de voz), transiciones de estado, cadenas y plantillas de respuesta, información de localización y cualquier otra información introducida a través de RADE para proporcionar la representación visual/declarativa de las funcionalidades de agente reactivo. La especificación de definición de agente puede implementarse dentro de una aplicación activada por voz (por ejemplo, un asistente digital personal nativo del sistema operativo del dispositivo o una aplicación activada por voz de tercera parte) junto con uno o más clasificadores de entendimiento de lenguaje (una definición del término "clasificador" se proporciona aquí más adelante). Cada clasificador también puede estar asociado con uno o más de un dominio, intento y ranura, así como con una expresión de usuario.

En los casos cuando una expresión de usuario (o entrada de texto) no coincide con una expresión/comando específico dentro de la especificación de definición de agente, una herramienta de etiquetado de usuario final (EULT) puede ser utilizada en el dispositivo de cómputo para permitir que el usuario seleccione uno o más de un dominio, intento para el dominio y/o una o más ranuras para el intento. En casos en que el dominio no esté disponible, el usuario puede agregar un dominio y, opcionalmente, especificar un intento y/o ranura para ese dominio. Un conjunto de datos etiquetados puede ser creado al asociar la expresión de usuario con el dominio seleccionado, intento y/o ranura. Un clasificador asociado con el intento seleccionado (y/o dominio o ranura) entonces puede

ser actualizado mediante el conjunto de datos etiquetados. La actualización para el clasificador puede activarse sólo después de que un determinado número de usuarios hace una selección de usuario sustancialmente similar (es decir, solicitando el mismo dominio o uno similar, intento y/o ranura), para evitar la manipulación y actualización fraudulenta de un clasificador. La actualización para el clasificador puede hacerse localmente (dentro del dispositivo de cómputo) y el clasificador actualizado puede ser luego almacenado en una base de datos de nube, en donde puede ser utilizado por otros usuarios. Alternativamente, la información de selección de usuario puede ser enviada a una computadora de servidor (servidor de la nube) en donde el conjunto de datos etiquetados puede ser creado y el clasificador actualizado después de que un número suficiente de usuarios realicen la misma expresión (o similar) y selección de usuario.

En este documento, se detallan diversos métodos, procesos y procedimientos. Aunque pasos particulares pueden ser descritos en cierta secuencia, dicha secuencia es principalmente por conveniencia y claridad. Un paso particular puede repetirse más de una vez, puede ocurrir antes o después de otros pasos (incluso si esos pasos se describen de otra manera en otra secuencia) y puede ocurrir en paralelo con otros pasos. Un segundo paso es necesario para seguir un primer paso sólo cuando el primer paso deba ser completado antes de que el segundo paso comience. Tal situación será específicamente señalada cuando no se desprende claramente

del contexto. Un paso particular puede ser omitido; un paso particular es requerido sólo cuando su omisión podría tener un impacto material sobre otro paso.

En este documento, se utilizan los términos "y", "o" e "y/o".

5 Tales términos son para ser leídos con el mismo significado; es decir, inclusive. Por ejemplo, "A" y "B" puede significar, al menos, lo siguiente: "tanto A como B", "sólo A", "sólo B", "al menos tanto A como B". Como otro ejemplo, "A" o "B" puede significar, al menos, lo siguiente: "sólo A", "sólo B", "tanto A como B", "al menos tanto A
10 como B". Cuando un exclusivo o se pretende, esto será específicamente señalado (por ejemplo, "A o B", "a lo mucho uno de A y B").

En este documento, se describen varios métodos implementados por computadora, procesos y procedimientos. Se
15 entiende que las varias acciones (recepción, almacenamiento, envío, comunicación, visualización, etc.) son realizadas por un dispositivo de *hardware*, incluso si la acción puede ser autorizada, iniciada o activada por un usuario o, incluso, si el dispositivo de *hardware* es controlado por un programa de computadora, *software*, *firmware*, etc.

20 Además, se entiende que el dispositivo de *hardware* está operando en datos, incluso si los datos pueden representar conceptos u objetos del mundo real, por lo que el etiquetado explícito como "datos" como tal se omite. Por ejemplo, cuando el dispositivo de *hardware* es descrito como "almacenar un registro", se entiende que
25 el dispositivo de *hardware* está almacenando datos que representa el

registro.

Tal como se utiliza aquí, el término "agente" o "agente reactivo" se refiere a una estructura de datos/comando que puede ser utilizada por un asistente digital personal para implementar uno o
5 más diálogos de respuesta (por ejemplo, voz, texto y/o respuestas táctiles) asociados con una funcionalidad del dispositivo. La funcionalidad del dispositivo (por ejemplo, correo electrónico, mensajería, etc.) puede ser activada por una entrada de usuario (por ejemplo, el comando de voz) al asistente digital personal. El agente
10 reactivo (o agente) puede definirse utilizando un documento de XML de definición de agente de voz (VAD), definición de comando de voz (VCD), o definición de agente reactivo (RAD) (o cualquier otro tipo de documento legible por computadora), así como un código de programación (por ejemplo, código C++) utilizado para impulsar al
15 agente a través del diálogo. Por ejemplo, se puede utilizar un agente reactivo de correo electrónico para, basado en tareas del usuario (por ejemplo, comandos de voz), abrir una nueva ventana de correo electrónico, redactar un correo electrónico basado en la entrada de voz y enviar el correo electrónico a una dirección de correo
20 electrónico especificada una entrada de voz a un asistente digital personal. Un agente reactivo también puede ser usado para proporcionar una o más respuestas (por ejemplo, respuestas de audio/vídeo/táctiles) durante una sesión de diálogo iniciada con un asistente digital personal basándose en la entrada de usuario.

25 Tal como se utiliza aquí, el término "esquema XML" se refiere a

un documento con un conjunto de segmentos de código XML que se utilizan para describir y validar datos en un entorno de XML. Más específicamente, el esquema XML puede listar elementos y atributos que se utilizan para describir el contenido en un documento XML, en
5 donde cada elemento está permitido, qué tipo de contenido está permitido, y así sucesivamente. Un usuario puede generar un archivo XML (por ejemplo, para usarse en una definición de agente reactivo), que se adhiere al esquema XML.

Tal como se utiliza aquí, el término "dominio" puede ser
10 utilizado para indicar un campo o rango de conocimiento personal y puede estar asociado con una categoría de funciones realizadas por un dispositivo de cómputo. Los dominios ilustrativos incluyen correo electrónico (por ejemplo, un agente de correo electrónico puede ser utilizado por un asistente digital personal (DPA) para generar/enviar
15 un correo electrónico), mensaje (por ejemplo, un agente de mensaje que puede ser utilizado por un DPA para generar/enviar mensajes de texto), alarma (un agente reactivo de alarma puede utilizarse para configurar/eliminar/modificar alarmas) y así sucesivamente.

Tal como se utiliza aquí, el término "intento" puede ser
20 utilizado para indicar al menos una acción utilizada para realizar al menos una función de la categoría de funciones para un dominio identificado. Por ejemplo, un intento de "configurar una alarma" puede ser utilizado para un dominio de alarma.

Tal como se utiliza aquí, el término "ranura" puede ser utilizado
25 para indicar un valor específico o un conjunto de valores que se

utilizan para completar una acción específica para un par de intento-dominio dado. Una ranura puede estar asociada a uno o más intentos y puede ser explícitamente proporcionada (es decir, anotada) en la plantilla del esquema XML. Típicamente, el dominio, intento y una o
5 más ranuras hacen una construcción de entendimiento de lenguaje, sin embargo dentro de un escenario agente dado, una ranura podría compartirse entre varios intentos. Como un ejemplo, si el dominio es una alarma con dos diferentes intentos - establecer una alarma y eliminar una alarma, entonces ambos estos intentos podrían
10 compartir la misma ranura "Tiempoalarma". En este sentido, una ranura puede estar conectada a uno o más intentos.

Tal como se utiliza aquí, el término "selección de usuario" (en relación con la herramienta de etiquetado de usuario final) se refiere a una selección por parte del usuario de información de dominio y/o
15 intento y/o ranura. En este sentido, una selección individual de un dominio o una intento o una ranura es posible (por ejemplo, sólo el intento puede ser seleccionado), así como cualesquiera pares (por ejemplo, selección de dominio-intento y ninguna ranura).

Tal como se utiliza aquí, el término "clasificador" o "clasificador
20 de entendimiento de lenguaje" se refiere a un algoritmo estadístico, basado en reglas o basado en aprendizaje por máquina o implementación de *software* que puede asignar una entrada de usuario dada (voz o texto) a un dominio e intento. También el algoritmo podría dar salida a una puntuación de confianza para
25 cualquier clasificación que se realiza utilizando el clasificador. El

mismo algoritmo o una pieza subsecuente de *software* puede entonces inferir/determinar el conjunto de ranuras especificadas por el usuario como parte de la expresión de ese par de dominio-intento. Una expresión de usuario dada puede entrenar múltiples
5 clasificadores - algunos para el caso positivo y otros para el caso negativo. Como un ejemplo, una expresión de usuario (o un comando de voz/texto) "mensaje de Rob estoy corriendo tarde" podría utilizarse para entrenar un clasificador de "Mensajería" como un conjunto de entrenamiento positivo y el clasificador de "correo
10 electrónico" como un conjunto de entrenamiento negativo. Un clasificador puede estar asociado con una o más partes de datos etiquetados (por ejemplo, la expresión del usuario, dominio, intento y/o ranura).

La Figura 1 es un diagrama de bloques que ilustra un ejemplo
15 de una arquitectura (100) para la actualización de modelos de clasificador de entendimiento de lenguaje, de acuerdo con una modalidad ilustrativa de la descripción. Refiriéndose a la Figura 1, un dispositivo de cómputo de cliente (por ejemplo, teléfono inteligente u otro dispositivo de cómputo móvil tal como el dispositivo 800 en la
20 Figura 8) puede ejecutar *software* organizado según la arquitectura 100 para proporcionar la actualización de modelos de clasificador de entendimiento de lenguaje.

La arquitectura 100 incluye un dispositivo de cómputo 102 (por ejemplo, un teléfono, tableta, portátil, de escritorio u otro tipo de
25 dispositivo de cómputo) acoplado a una computadora de servidor

remoto (o computadoras) 140 a través de la red 130. El dispositivo de cómputo 102 incluye un micrófono 106 para convertir el sonido en una señal eléctrica. El micrófono 106 puede ser un micrófono dinámico, de condensador o piezoeléctrico que utiliza inducción
5 electromagnética, un cambio en la capacitancia o piezoelectricidad, respectivamente, para producir la señal eléctrica desde variaciones de presión de aire. El micrófono 106 puede incluir un amplificador, uno o más filtros analógicos o digitales y/o un convertidor analógico-a-digital para producir una entrada de sonido digital. La entrada de
10 sonido digital puede comprender una reproducción de la voz del usuario, como, por ejemplo, cuando el usuario solicita al asistente digital personal 110 a realizar una tarea.

El asistente digital personal 110 se ejecuta en el dispositivo de cómputo 102 y permite que el usuario del dispositivo de cómputo 102
15 realice diversas acciones utilizando entrada de voz (o texto). El asistente digital personal 110 puede comprender un módulo de procesamiento de lenguaje natural 112, una estructura de definición de agente 114, interfaces de usuario 116, modelo de clasificador de entendimiento de lenguaje (LUCM) 120 y una herramienta de
20 etiquetado de usuario final (EULT) 118. El asistente digital personal 110 puede recibir la entrada de voz de usuario mediante el micrófono 106, determinar una tarea correspondiente (por ejemplo, un comando de voz) de la entrada de voz de usuario usando la estructura de definición de agente 114 (por ejemplo, una estructura de datos de
25 comando de voz o una estructura de definición de agente reactivo) y

realizar la tarea (por ejemplo, comando de voz). En algunas situaciones, el asistente digital personal 110 envía el comando (voz o texto) de usuario a una de las aplicaciones activadas por voz de tercera parte 108. En otras situaciones, el asistente digital personal 110 controla la tarea en sí.

El sistema operativo de dispositivo (OS) 104 maneja las funciones de entrada de usuario, funciones de salida, funciones de acceso al almacenamiento, funciones de comunicación en red, y otras funciones para el dispositivo 110. El OS de dispositivo 104 proporciona acceso a las funciones del asistente digital personal 110.

La estructura de definición de agente 114 puede definir uno o más agentes del DPA 110 y puede especificar las tareas o comandos (por ejemplo, comandos de voz) soportados por el DPA 110 y/o las aplicaciones activadas por voz de tercera parte 108 junto con variaciones de comando de voz asociadas y ejemplos de comandos de voz. En algunas implementaciones, la estructura de definición de agente 114 se implementa en un formato XML. Además, la estructura de definición de agente 114 puede identificar aplicaciones activadas por voz disponibles de forma remota desde un almacenamiento de aplicación 146 y/o servicios activados por voz disponibles de forma remota desde un servicio web 148 (por ejemplo, accediendo a una definición de esquema disponible de computadoras de servidor remotas 140 que definen las capacidades para las aplicaciones remotas y/o los servicios remotos).

La estructura de definición de agente 114 puede proporcionarse junto con el modelo de clasificador de entendimiento de lenguaje (LUCM) 120 (por ejemplo, como parte del sistema operativo 104 o puede ser instalado en el momento que el DPA 110 se instala). El LUCM 120 puede incluir una pluralidad de clasificadores $C_1, \dots, C_n$, en donde cada clasificador puede estar asociado con uno o más de un dominio ($D_1, \dots, D_n$), intento ($I_1, \dots, I_n$) y/o una ranura ($S_1, \dots, S_n$). Cada uno de los clasificadores puede incluir un algoritmo estadístico, basado en reglas o basado en aprendizaje por máquina o implementación de *software* que puede asignar una entrada de usuario dada (voz o texto) a un dominio e intento. También el algoritmo podría dar salida a una puntuación de confianza para cualquier clasificación se está realizando utilizando el clasificador. En algunas implementaciones, un clasificador puede estar asociado con una o más de una información de dominio, intento y/o ranura y puede proporcionar una puntuación de confianza cuando se aplica a una entrada de voz o texto de usuario dada (por ejemplo, el escenario de implementación se describe con referencia a la Figura 2).

Aunque el LUCM 120 se ilustra como parte del DPA 110 junto con la estructura de definición de agente 114, la presente descripción no se limita a este respecto. En algunas modalidades, el LUCM 120 puede ser una copia local de un modelo de clasificador, que incluye los clasificadores ($C_1, \dots, C_n$) que son relevantes a la estructura de definición de agente 114 y el DPA 110. Otro modelo de

clasificador (por ejemplo, global) (por ejemplo, LUCM 170) puede ser almacenado en la nube (por ejemplo, como parte de las computadoras de servidor 140). El LUCM global 170 puede ser utilizado en el momento en que una estructura de definición de agente se crea de manera que un subconjunto de clasificadores (por ejemplo, relevantes) puede incluirse en dicha estructura de definición e implementarse como parte de una aplicación (por ejemplo, aplicación de tercera parte 108, el DPA 110 y/o el OS 104).

El DPA 110 puede procesar la entrada de voz de usuario utilizando un módulo de procesamiento de lenguaje natural 112. El módulo de procesamiento de lenguaje natural 112 puede recibir la entrada de sonido digital y traducir palabras pronunciadas por un usuario en texto usando el reconocimiento de voz. El texto extraído puede ser semánticamente analizado para determinar una tarea (por ejemplo, un comando de voz de usuario). Mediante el análisis de la entrada de sonido digital y la toma de acciones en respuesta a comandos hablados, el asistente digital personal 110 puede ser controlado por la entrada de voz del usuario. Por ejemplo, el asistente digital personal 110 puede comparar el texto extraído con una lista de posibles comandos de usuario potenciales (por ejemplo, almacenados en la estructura de definición de agente 114) para determinar el comando que más probablemente coincide con el intento del usuario. El DPA 110 también puede aplicar uno o más de los clasificadores del LUCM 120 para determinar una puntuación de confianza, seleccionar un clasificador basado en la puntuación de

confianza y determinar que un comando más probablemente coincida con el intento del usuario basado en el comando (o expresión) asociado con el clasificador. En este sentido, la coincidencia puede basarse en métodos estadísticos o probabilísticos, árboles de decisión u otras reglas, otros criterios de coincidencia adecuados o combinaciones de los mismos. Los comandos de usuario potenciales pueden ser comandos nativos del DPA 110 y/o comandos definidos en la estructura de definición de agente 114. Así, mediante la definición de comandos en la estructura de definición de agente 114 y los clasificadores dentro del LUCM 120, la gama de tareas que se pueden realizar en nombre del usuario por el DPA 110 puede ser extendida. Los comandos potenciales también pueden incluir comandos de voz para realizar tareas de las aplicaciones activadas por voz de tercera parte 108.

El asistente digital personal 110 incluye interfaces de voz y/o gráficas de usuario 116. Las interfaces de usuario 116 pueden proporcionar información al usuario describiendo las capacidades del DPA 110 (por ejemplo, capacidades de la EULT 118) y/o las aplicaciones activadas por voz de tercera parte 108.

La herramienta de etiquetado de usuario final (EULT) 118 puede comprender una lógica adecuada, circuitería, interfaces y/o código y puede ser operable para proporcionar funcionalidades para actualizar modelos de clasificador de entendimiento de lenguaje, tal como se describe aquí. Por ejemplo, la EULT 118 puede activarse en los casos cuando la estructura de definición de agente 114 no tenga

una cadena de comandos de voz que coincida con el comando de voz/texto del usuario o uno o más de los clasificadores disponibles devuelvan una puntuación de confianza que está por debajo de una cantidad de umbral (como se ve en la Figura 2). El usuario entonces

5 puede utilizar la EULT 118 para seleccionar un dominio, intento y/o ranura y asociar una tarea (por ejemplo, un comando de voz expresada como expresión) o comando de texto con la información de dominio, intento y/o ranura seleccionada por el usuario. Las selecciones del usuario y el comando de voz/texto introducido por el

10 usuario puede ser enviado a las computadoras de servidor 140, en donde el conjunto de clasificador global 170 puede ser actualizado (por ejemplo, un clasificador que coincide con el comando de voz/texto de usuario se actualiza con el dominio introducido por el usuario, intento y/o ranura). En este sentido, se puede utilizar el

15 aspecto de externalización masiva para entrenar/etiquetar clasificadores y, por lo tanto, mejorar el LUCM global y local (170 y 120).

El asistente digital personal 110 puede acceder a servicios remotos 142 que se ejecutan en computadoras de servidor remotas

20 140. Los servicios remotos 142 pueden incluir funciones de *software* provistas en una dirección de red a través de una red, tal como una red 130. La red 130 puede incluir una red de área local (LAN), una red de área amplia (WAN) o Internet, una intranet, una red cableada, una red inalámbrica, una red celular, combinaciones de los mismos,

25 o cualquier red adecuada para proporcionar un canal de

comunicación entre el dispositivo de cómputo 102 y las computadoras de servidor remotas 140. Cabe destacar que la topología de la red que se ilustra en la Figura 1 se ha simplificado y que múltiples redes y dispositivos de red se pueden utilizar para interconectar los varios sistemas de cómputo aquí descritos.

Los servicios remotos 142 pueden incluir varios servicios de cómputo que son accesibles desde las computadoras de servidor remotas 140 a través de la red 130. Los servicios remotos 142 pueden incluir un servicio de procesamiento de lenguaje natural 144 (por ejemplo, llamado por el asistente digital personal 110 para realizar o ayudar con funciones del procesamiento de lenguaje natural del módulo 112). Los servicios remotos 142 pueden incluir un almacenamiento de aplicación 146 (por ejemplo, un almacenamiento de aplicación ofreciendo aplicaciones activadas por voz que puede buscarse o descargarse e instalarse). Los servicios remotos 142 también pueden incluir servicios web 148, los cuales pueden accederse a través de entrada de voz utilizando el asistente digital personal 110. Los servicios remotos 142 también pueden incluir una herramienta de etiquetado de programador 150, un servicio de capacitación de modelo de clasificador 152 y un servicio de detección de fraude de modelo de clasificador 154, tal como se explica a continuación. Las computadoras de servidor remotas 140 también pueden manejar una base de datos de expresiones 160 y una base de datos etiquetados 162.

La Figura 2 es un diagrama de bloques 200 que ilustra varios

usos de clasificadores de entendimiento de lenguaje por aplicaciones activadas por voz, de acuerdo con una modalidad ilustrativa de la descripción. Refiriéndose a las Figuras 1-2, un usuario (por ejemplo, usuario del dispositivo 102) puede introducir una entrada de voz 202.

5 El bloque de reconocimiento de voz 206 (por ejemplo, 112) puede convertir la voz de entrada 202 en un comando de usuario (texto) 208. El comando de usuario 208 puede, alternativamente, ser introducido como entrada de texto 204. En el bloque 210, se puede realizar una coincidencia de definición de agente al hacer coincidir el

10 comando de usuario 208 con uno o más comandos de usuario especificados en la estructura de definición de agente (por ejemplo, 114). Si hay una coincidencia directa (en 212), entonces el dominio 216, el intento 218 y/o la ranura 220 puede inferirse del comando de usuario coincidente y dicha información puede ser utilizada por el

15 DPA 110 y/o la aplicación 108) en el bloque 232. Sin embargo, si no hay ninguna coincidencia (en 214), entonces la coincidencia utilizando el LUCM 120 (o 170) puede ser realizada.

Más específicamente, el comando de usuario 208 se puede utilizar como entrada en los clasificadores C_1 , ..., C_n y las

20 puntuaciones de confianza correspondientes 240 pueden ser calculadas. Si para un determinado clasificador (por ejemplo, C_1), la puntuación de confianza es mayor o igual que un valor de umbral (por ejemplo, 20%), entonces el clasificador puede usarse para extraer el dominio 224, el intento 226, y/o la ranura 228 asociados

25 con dicho clasificador. El dominio/intento/slot extraídos puede ser

utilizado por el DPA 110 o la aplicación 108 (en 230). Si la puntuación de confianza, sin embargo, es menor que el umbral (por ejemplo, en 250), entonces el modelo de clasificador puede ser actualizado (por ejemplo, utilizando la EULT 118 y como se ve en las Figuras 3B-4B). El dominio, intento y/o ranura determinada durante el proceso de etiquetado de EULT puede ser utilizado por el DPA 110 y/o aplicación 108 (en 232).

Aunque se utiliza una puntuación de confianza generada por los clasificadores (junto con un valor de umbral) para determinar si se desea utilizar la EULT para obtener una selección de usuario, la presente descripción no se limita a este respecto. En otra implementación, se pueden utilizar otros métodos (por ejemplo, en lugar de utilizar un valor de umbral) para determinar si se va a utilizar la EULT para recibir una selección de usuario de al menos uno de una información de dominio, intento y/o ranura.

Las Figuras 3A-3B muestran ciclos de procesamiento ilustrativos para actualizar los modelos de clasificador de entendimiento de lenguaje, de acuerdo con una modalidad ilustrativa de la descripción. Refiriéndose a la Figura 3A, se ilustra la arquitectura 300 capacitación/actualización de los datos de clasificador utilizando una herramienta de etiquetado de programador 150. Como se ve en la Figura 3A, una estructura de definición de agente 114 puede estar integrada con el LUCM 120 (el LUCM 120 puede ser el mismo, o un subconjunto del LUCM 170). La estructura de definición de agente 114 y el LUCM 120 entonces pueden ser

implementados como parte de la aplicación 108 (por ejemplo, disponible en el almacenamiento de aplicación 146) o el DPA 110. La aplicación 108 (y el DPA 110) puede ser instalada en el dispositivo 102.

5 En los casos en que la EULT 118 esté desactivada, el usuario puede proporcionar una expresión 302 (por ejemplo, el comando de usuario). La expresión puede ser comunicada y almacenada como parte de la base de datos de expresiones 160, que también puede almacenar expresiones de los usuarios de otros dispositivos de
10 cómputo comunicativamente acoplados a las computadoras de servidor 140. Un administrador/desarrollador de red puede, entonces, utilizar la herramienta de etiquetado desarrollador 150 para recuperar una expresión (por ejemplo, 302) de la base de datos 160 y generó una selección de dominio, intento y/o ranura, 303. La
15 selección de administrador 303 puede combinarse con la expresión 302 y almacenarse como datos etiquetados dentro de la base de datos etiquetados 162. El administrador puede entonces pasar los datos etiquetados a lo largo del servicio de capacitación de clasificador 152 (o los datos etiquetados automáticamente pueden
20 ser comunicados al servicio de capacitación 152 después de ser almacenados en la base de datos 162).

 El servicio de capacitación de modelo de clasificador 152 puede incluir una lógica adecuada, circuitería, interfaces, y/o código y puede ser operable para realizar la capacitación (o actualización)
25 de uno o más clasificadores dentro de los LUCM 120 y/o 170.

Durante la capacitación del clasificador 304 ilustrativo, el conjunto de datos etiquetados puede ser recuperado (por ejemplo, 302 y 303); se puede utilizar la información de dominio, intento y/o ranura (por ejemplo, 303) (por ejemplo, como un índice) para acceder al LUCM 120/170 y recuperar un clasificador que está asociado con ese dominio, intento y/o ranura. El servicio de capacitación 152 puede entonces actualizar el clasificador de modo que esté asociado con la expresión/comando de usuario (302), así como uno o más del dominio, intento y/o ranura (303) proporcionado por el administrador utilizando la herramienta de etiquetado de programador 150. El LUCM actualizado 120 entonces puede utilizarse y combinarse con una estructura de definición de agente para implementación en una aplicación.

Refiriéndose a la Figura 3B, se ilustra la arquitectura 370 para la capacitación/actualización de los datos de clasificador utilizando una herramienta de etiquetado de usuario final (EULT) 118. Como se ve en la Figura 3B, una estructura de definición de agente 114 puede combinarse con el LUCM 120 (el LUCM 120 puede ser el mismo, o un subconjunto del LUCM 170). La estructura de definición de agente 114 y el LUCM 120 entonces pueden ser implementados como parte de la aplicación 108 (por ejemplo, como disponible en el almacenamiento de aplicación 146), el DPA 110 y/o aplicaciones 350, ..., 360. Las aplicaciones 108, 350, ..., 360 (y el DPA 110) pueden ser instaladas en el dispositivo 102.

En los casos en que la EULT 118 es activada, un usuario

puede proporcionar una expresión 302 (por ejemplo, comando de usuario). La expresión se puede comunicar y almacenar como parte de la base de datos de expresiones 160, que también puede almacenar las expresiones de los usuarios de otros dispositivos de cómputo comunicativamente acoplados a las computadoras de servidor 140. El usuario del dispositivo 102 puede entonces utilizar la EULT 118 para proporcionar la entrada de usuario, selecciona uno o más de un dominio, intento y/o ranura asociada con la expresión/comando 302 (esto es suponiendo que no hay una coincidencia directa (por ejemplo, 212) con un comando dentro de la estructura de definición de agente 114 y no hay ninguna puntuación de confianza que esté por encima de un valor umbral (por ejemplo, 240)).

El usuario puede utilizar la EULT 118 para seleccionar un dominio, intento y/o ranura (por ejemplo, 320) asociados con la expresión 302. El DPA 110 (o de lo contrario el dispositivo 102) puede seleccionar al menos uno de los clasificadores Cl , ..., Cn en el LUCM 120 como coincidente con la selección de usuario introducida 320 (por ejemplo, un clasificador puede ser seleccionado del LUCM 120 basándose en la información de coincidencia de dominio, intento y/o ranura asociada con el clasificador con la información de dominio, intento, y/o ranura de la selección de usuario 320 introducida a través de la EULT 118).

De acuerdo con una modalidad ilustrativa de la descripción, después de un clasificador de coincidencia se recupera del LUCM

120, el dispositivo 102 puede actualizar el clasificador (por ejemplo, como se discutió anteriormente en referencia a 304) y almacenar el clasificador actualizado/capacitado como un clasificador local 330. La capacitación y actualización del clasificador y la generación del clasificador local 330 pueden realizarse utilizando el servicio de capacitación de modelo de clasificador 152 de computadoras de servidor remotas 140. En este sentido, uno o más de los clasificadores locales 330 pueden ser generados, sin tales clasificadores capacitados presentes en el LUCM global 170. Los clasificadores locales 330 pueden estar asociados con un perfil de usuario 340 y pueden ser utilizados/compartidos entre una o más de las aplicaciones 350, ..., 360 instaladas en el dispositivo 102. Opcionalmente, los clasificadores locales 330 pueden ser almacenados en las computadoras de servidor 140, como parte del perfil de usuario 340 (un perfil también puede ser almacenado en las computadoras de servidor 140, junto con otra información de cuenta de usuario/perfil).

El DPA 110 también puede comunicar la información de dominio, intento y/o ranura 320 seleccionado por el usuario, junto con la expresión 302, para almacenamiento como datos etiquetados dentro de la base de datos de datos etiquetados 162. Los datos etiquetados pueden ser entonces pasados al servicio de capacitación de clasificador 152 para la capacitación. De acuerdo con una modalidad ilustrativa de la descripción, un servicio de detección de fraude de modelo de clasificador 154 puede ser utilizado en conexión

con el servicio de capacitación 152. Más concretamente, el servicio de detección de fraude 154 puede comprender una lógica adecuada, circuitería, interfaces, y/o código y puede ser operable para prevenir la capacitación/actualización de clasificador a menos que un cierto
5 número mínimo (umbral) de los usuarios han solicitado la misma (o sustancialmente similar) actualización a un clasificador asociado con la misma (o sustancialmente similar) pronunciación de usuario. En este sentido, se puede evitar una actualización automática de clasificador en casos cuando un usuario intenta asociar una tarea
10 (por ejemplo, una expresión para expresar un comando de voz) con un dominio, intento, y/o ranura que la mayoría de los demás usuarios en el sistema no asocia con dicha expresión.

Suponiendo que un número mínimo de usuarios ha solicitado la misma actualización o sustancialmente similar a un clasificador,
15 entonces la capacitación/actualización (304) del clasificador puede proceder, como se discutió anteriormente en referencia a la Figura 3A. Durante la capacitación de clasificador 304 ilustrativa, el conjunto de datos etiquetados puede ser recuperado (por ejemplo, 302 y 303); la información de dominio, intento y/o ranura (por
20 ejemplo, 303) puede utilizarse (por ejemplo, como un índice) para acceder al LUCM 120/170 y recuperar un clasificador que está asociado con ese dominio, intento y/o ranura. El servicio de capacitación 152 puede entonces actualizar el clasificador de modo que quede asociado con la expresión/comando de usuario (302), así
25 como uno o más del dominio, intento y/o ranura (303) proporcionado

por el administrador utilizando la herramienta de etiquetado de programador 150. El LUCM 120 actualizado puede utilizarse y combinarse con una estructura de definición de agente para implementación en una aplicación.

5 Las Figuras 4A-4B ilustran ejemplos de interfaces de usuario de una herramienta de etiquetado de usuario final, que se pueden utilizar de acuerdo con una modalidad ilustrativa de la descripción. Refiriéndose a la Figura 4A, la interfaz de usuario en 402 ilustra una vista inicial de un DPA 110 que solicita al usuario que proporcione
10 una tarea (por ejemplo, un comando de voz). En 404, el usuario proporciona un comando de voz en 405. En 406, el DPA 110 puede haber realizado el procesamiento (por ejemplo, 202-214) y puede haber determinado que no hay ningún comando de usuario coincidente en la estructura de definición de agente 114 o una
15 puntuación de confianza suficientemente alta (240). A continuación, el proceso continúa (por ejemplo, en 250) activando la interfaz de EULT 118. En 407, el DPA 110 notifica al usuario que la tarea (por ejemplo, comando de voz) es ambiguo y pregunta si el usuario desearía activar la "Herramienta de Etiquetado" (EULT 118). A
20 continuación, el usuario activa la EULT 118 pulsando el botón de *software* 408.

Refiriéndose a la Figura 4B, la interfaz de usuario en 409 sugiere uno o más dominios de modo que el usuario pueda seleccionar un dominio pertinente para su tarea (por ejemplo,
25 comando de voz). Uno o más dominios pueden ser listados (por

ejemplo, uno o más dominios relevantes (por ejemplo, fonéticamente similares) para la tarea (o comando de voz) o todos los dominios disponibles en el sistema). Después de que el usuario selecciona un dominio, la interfaz de usuario 410 puede utilizarse para enlistar uno
5 o más intentos asociados con el dominio seleccionado. Alternativamente, todos los intentos disponibles pueden ser enlistados para que el usuario pueda elegir. Después de que el usuario selecciona un intento, la interfaz de usuario 412 puede utilizarse para enlistar una o más ranuras asociadas con el intento
10 seleccionado. De forma alternativa, todas las ranuras disponibles pueden ser enlistadas para que el usuario pueda elegir. Después de seleccionar la ranura, la información del dominio, intento, y/o ranura 320 puede ser procesada como se ha descrito anteriormente.

Las Figuras 5-7 son diagramas de flujo que ilustran la
15 generación de una definición de agente reactivo, de acuerdo con una o más modalidades. Refiriéndose a las Figuras 1-5, el método ilustrativo 500 puede empezar en 502, cuando una primera selección de usuario (320) de al menos uno de los siguientes: al menos un intento de una pluralidad de intentos disponibles y/o al menos una
20 ranura para al menos una intento puede ser recibido de al menos un dispositivo de cómputo (por ejemplo, 102) de una pluralidad de dispositivos de cómputo comunicativamente acoplados a una computadora de servidor (por ejemplo, 140). El al menos un intento (intento en la selección de usuario 320) está asociado con al menos
25 una acción utilizada para realizar al menos una función de una

categoría de funciones para un dominio. La al menos una ranura (por ejemplo, dentro de la selección de usuario 320) indica un valor utilizado para realizar la al menos una acción. La primera selección de usuario (320) está asociada con una entrada de voz digital (por ejemplo, expresión 302) recibida en al menos un dispositivo de cómputo (102). En 504, tras la recepción de al menos otro dispositivo de cómputo de la pluralidad de dispositivos de cómputo, una pluralidad de selecciones de usuario subsiguientes que son idénticas a la primera selección de usuario, un conjunto de datos etiquetados se genera al emparejar la entrada de voz digital con la primera selección de usuario. Por ejemplo, después de 302 y 320 son emparejados para generar el conjunto de datos etiquetados, el servicio de capacitación 152 puede proseguir con la capacitación del clasificador correspondiente después de un cierto número (umbral) de otros usuarios presenta la misma (o sustancialmente similar) selección de usuario y expresión. En 506, el servicio de capacitación de modelo de clasificador 152 puede seleccionar un clasificador de entendimiento de lenguaje de una pluralidad de clasificadores de entendimiento de lenguaje disponibles (por ejemplo, del LUCM 170) asociados con una o más definiciones de agente. La selección puede basarse al menos en al menos un intento. En 508, el servicio de capacitación 152 puede actualizar el clasificador de entendimiento de lenguaje seleccionado basándose en el conjunto de datos etiquetados generado.

Refiriéndose a las Figuras 1-3B y 6, el método ilustrativo 600

puede empezar en 602, cuando una entrada de voz digital (302) de un usuario del dispositivo de cómputo (102) puede ser recibida a través de uno o más micrófonos (106) de un dispositivo de cómputo (102). En 604, el módulo de procesamiento de lenguaje natural 112
5 puede realizar el procesamiento de lenguaje natural utilizando la entrada de voz digital para determinar una solicitud de voz de usuario.

En 606, al determinar que la solicitud de voz de usuario no coincide (por ejemplo, 214) en al menos uno de una pluralidad de
10 comandos de voz predefinidos en una definición de agente (por ejemplo, 114) de un asistente digital personal (110) que se ejecuta en el dispositivo de cómputo, una selección de un usuario (320) de al menos uno de los siguientes: un intento de una pluralidad de intentos disponibles y al menos una ranura para al menos un intento puede
15 recibirse utilizando una interfaz gráfica de usuario de una herramienta de etiquetado de usuario final (EULT) (118) del dispositivo de cómputo (102). El intento está asociado con al menos una acción utilizada para realizar al menos una función de una categoría de funciones para un dominio y al menos una ranura que
20 indica un valor que se utiliza para realizar al menos una acción. En 608, el DPA 110 puede generar un conjunto de datos etiquetados al emparejar la solicitud de voz de usuario (320) y la selección de usuario (302). En 610, el DPA 110 (o dispositivo 102) puede seleccionar un clasificador de entendimiento de lenguaje de una
25 pluralidad de clasificadores de entendimiento de lenguaje disponibles

(por ejemplo, C1, ..., Cn en el LUCM 120) asociados con la definición de agente (por ejemplo, 114). La selección del clasificador puede basarse, al menos en al menos un intento seleccionado por el usuario mediante la EULT 118. En 612, el DPA 110 (o dispositivo 5 102) puede actualizar el clasificador de entendimiento de lenguaje seleccionado basado en el conjunto de datos etiquetados generados (por ejemplo, basado en 302 y 320, creando el clasificador local 330).

Refiriéndose a las Figuras 1-3B y 7, el método ilustrativo 700 10 puede empezar en 702, cuando una solicitud de usuario puede ser determinada basándose en la entrada de usuario (302) recibida en un dispositivo de cómputo (102). La solicitud de usuario se puede recibir a través de al menos uno de entrada de texto (204) y/o entrada de voz (202), la solicitud siendo para una funcionalidad de 15 un asistente digital personal (110) que se ejecuta en el dispositivo de cómputo. En 704, el DPA 110 (o dispositivo 102) puede determinar que la solicitud de usuario no coincide con al menos uno de una pluralidad de tareas predefinidas (por ejemplo, comandos de voz) en una definición de agente (114) del asistente digital personal (por 20 ejemplo, 214).

En 706, el DPA 110 (o dispositivo 102) puede generar una puntuación de confianza (240) mediante la aplicación de una pluralidad de clasificadores de entendimiento de lenguaje disponibles (C1, ..., Cn) asociados con la definición de agente para la solicitud 25 de usuario (208). En 708, tras determinar que la puntuación de

confianza es menor que un valor de umbral (250), el DPA 110 recibe mediante una interfaz gráfica de usuario de una herramienta de etiquetado de usuario final (EULT) (118) del dispositivo de cómputo, una selección de usuario (320) de al menos uno de los siguientes: al menos un intento de una pluralidad de intentos disponibles y al menos una ranura para al menos un intento. Al menos uno de los intentos está asociado con al menos una acción utilizada para realizar al menos una función de una categoría de funciones para un dominio y al menos una ranura que indica un valor que se utiliza para realizar al menos una acción.

En el 710, el DPA 110 (o dispositivo 102) genera un conjunto de datos etiquetados al emparejar la solicitud de voz de usuario (302) y la selección de usuario (320). En 712, el DPA 110 (o dispositivo 102) selecciona un clasificador de entendimiento de lenguaje de la pluralidad de clasificadores de entendimiento de lenguaje disponibles (LUCM 120) asociados con la definición de agente, la selección basada por lo menos en un intento seleccionado por el usuario. En 714, el DPA 110 (o dispositivo 102) genera un clasificador de entendimiento de lenguaje actualizado mediante la capacitación del clasificador de entendimiento de lenguaje seleccionado mediante el conjunto de datos etiquetados generados (por ejemplo, generar un clasificador local 330).

La Figura 8 es un diagrama de bloques que ilustra un ejemplo de dispositivos de cómputo móviles en conjunción con el cual las innovaciones descritas aquí pueden ser implementadas. El

dispositivo móvil 800 incluye una variedad de componentes de *hardware* y *software* opcionales, mostrados generalmente en 802. En general, un componente 802 en el dispositivo móvil se puede comunicar con cualquier otro componente del dispositivo, aunque no
5 todas las conexiones se muestran, para facilidad de ilustración. El dispositivo móvil 800 puede ser cualquiera de una variedad de dispositivos de cómputo (por ejemplo, teléfono celular, teléfono inteligente, computadora portátil, computadora laptop, portátil, dispositivo de tableta, netbook, reproductor multimedia, asistente
10 digital personal (PDA), cámara, cámara de vídeo, etc.) y puede permitir comunicaciones de dos vías inalámbricas con una o más redes de comunicaciones móviles 804, tales como Wi-Fi, celulares, o red de satélite.

El dispositivo móvil ilustrado 800 incluye un controlador o
15 procesador 810 (por ejemplo, procesador de señal, microprocesador, ASIC o cualquier otro circuito lógico de control y procesamiento) para realizar tareas tales como la codificación de señal, el procesamiento de datos (incluyendo la asignación de pesos y datos de clasificación tales como resultados de búsqueda), procesamiento
20 de entrada/salida, control de energía y/u otras funciones. Un sistema operativo 812 controla la asignación y uso de los componentes 802 y soporta uno o más programas de aplicación 811. El sistema operativo 812 puede incluir una herramienta de etiquetado de usuario final 813, la cual puede tener funcionalidades que son similares a las
25 funcionalidades de la EULT 118 descrita con referencia a las Figuras

1-7.

El dispositivo móvil ilustrado 800 incluye memoria 820. La memoria 820 puede incluir memoria no removible 822 y/o memoria removible 824. La memoria no removible 822 puede incluir RAM, ROM, memoria flash, un disco duro u otras tecnologías de almacenamiento en memoria bien conocidas. La memoria removible 824 puede incluir la memoria flash o una tarjeta módulo de identidad de suscriptor (SIM), que es bien conocida en el Sistema Global para sistemas de comunicación de Comunicaciones Móviles (GSM), u otras tecnologías de almacenamiento de memoria bien conocidas, tales como "tarjetas inteligentes". La memoria 820 puede utilizarse para almacenar y/o codificar datos para que se ejecute el sistema operativo 812 y las aplicaciones 811. Datos ilustrativos pueden incluir páginas web, texto, imágenes, archivos de sonido, datos de vídeo u otros conjuntos de datos para ser enviados a y/o recibidos desde uno o más servidores de red u otros dispositivos a través de una o más redes de cable o inalámbricas. La memoria 820 puede ser utilizada para almacenar un identificador de suscriptor, tal como una Identidad de suscriptor móvil internacional (IMSI) y un identificador de equipo, tales como un Identificador de equipos móviles internacional (IMEI). Estos identificadores pueden ser transmitidos a un servidor de red para identificar a los usuarios y equipos.

El dispositivo móvil 800 puede soportar uno o más dispositivos de entrada 830, tales como una pantalla táctil 832 (por ejemplo, capaz de capturar entradas de toque de dedo, entradas de gesto de

dedo o entradas de pulsaciones de teclas de un teclado virtual o teclado), micrófono 834 (por ejemplo, capaz de capturar entrada de voz), cámara 836 (por ejemplo, capaz de capturar imágenes fijas y/o imágenes de vídeo), teclado físico 838, botones y/o seguidor 840 y uno o más dispositivos de salida 850, como, por ejemplo, un altavoz 852 y una pantalla 854. Otros posibles dispositivos de salida (no mostrados) pueden incluir dispositivos de salida piezoeléctricos u otros hápticos. Algunos dispositivos pueden servir a más que una función de entrada/salida. Por ejemplo, la pantalla táctil 832 y pantalla 854 pueden ser combinadas en un solo dispositivo de entrada/salida. El dispositivo móvil 800 puede proporcionar una o más interfaces de usuario naturales (NUIs). Por ejemplo, el sistema operativo 812 o aplicaciones 811 pueden comprender *software* de procesamiento multimedia, tal como reproductor de audio/vídeo.

Un módem inalámbrico 860 puede ser acoplado a una o más antenas (no mostradas) y puede soportar la comunicación bidireccional entre el procesador 810 y dispositivos externos, como es bien entendido en la técnica. El módem 860 se muestra de forma genérica y puede incluir, por ejemplo, un módem celular para comunicarse a larga distancia con la red de comunicación móvil 804, un módem compatible con Bluetooth 864, o un módem compatible con Wi-Fi 862 para la comunicación a corta distancia con un dispositivo equipado con Bluetooth externo o una red de datos inalámbrica local o enrutador. El módem inalámbrico 860 normalmente está configurado para la comunicación con una o más

redes celulares, tales como una red GSM para comunicaciones de voz y datos en una sola red celular, entre redes celulares, o entre el dispositivo móvil y una red telefónica conmutada pública (PSTN).

El dispositivo móvil además puede incluir al menos un puerto
5 de entrada/salida 880, un suministro de energía 882, un receptor de sistema de navegación por satélite 884, tal como un receptor de sistema de posicionamiento global (GPS), sensores 886 tales como un acelerómetro, un giroscopio o un sensor de proximidad por infrarrojos para detectar el movimiento y orientación del dispositivo
10 800 y para recibir los comandos de gestos como entrada, un transceptor 888 (para transmitir de forma inalámbrica señales analógicas o digitales) y/o un conector físico 890, que puede ser un puerto USB, un puerto IEEE 1394 (FireWire), y/o un puerto RS-232. Los componentes ilustrados 802 no son necesarios o todos
15 inclusivos, ya que cualquiera de los componentes mostrados puede ser borrado y otros componentes pueden ser añadidos.

El dispositivo móvil puede determinar datos de ubicación que indican la ubicación del dispositivo móvil basándose en la información recibida a través del receptor de sistema de navegación
20 por satélite 884 (por ejemplo, receptor de GPS). Alternativamente, el dispositivo móvil puede determinar los datos de ubicación que indican la ubicación del dispositivo móvil de otra manera. Por ejemplo, la ubicación del dispositivo móvil puede ser determinada mediante triangulación entre torres celulares de una red celular. O
25 bien, la ubicación del dispositivo móvil puede determinarse en base a

las ubicaciones conocidas de enrutadores de Wi-Fi en las proximidades del dispositivo móvil. Los datos de ubicación se pueden actualizar cada segundo o en alguna otra base, dependiendo de la implementación y/o la configuración de usuario. Independientemente
5 de la fuente de datos de ubicación, el dispositivo móvil puede proporcionar los datos de ubicación para asignar una herramienta de navegación útil para usar en la navegación de mapa.

Como un dispositivo de cómputo de cliente, el dispositivo móvil 800 puede enviar peticiones a un dispositivo de cómputo de servidor
10 (por ejemplo, un servidor de búsqueda, un servidor de rutas, etc.) y recibir imágenes de mapa, distancias, direcciones y otros datos de mapa, resultados de búsqueda (por ejemplo, basado en POI en una búsqueda POI dentro de un área de búsqueda designada) u otros datos de retorno desde el dispositivo de cómputo de servidor.

15 El dispositivo móvil 800 puede ser parte de un entorno de implementación en el que diversos tipos de servicios (por ejemplo, servicios de cómputo) son proporcionados por una "nube" de cómputo. Por ejemplo, la nube puede comprender una colección de dispositivos de cómputo, los cuales pueden ser ubicados de forma
20 centralizada o distribuida, que proporcionan servicios basados en la nube para diferentes tipos de usuarios y dispositivos conectados a través de una red, tal como Internet. Algunas tareas (por ejemplo, entrada de procesamiento de usuario y presentación de una interfaz de usuario) pueden realizarse en los dispositivos de cómputo locales
25 (por ejemplo, dispositivos conectados) mientras que otras tareas (por

ejemplo, el almacenamiento de datos que serán utilizados en el procesamiento posterior, carga de datos y clasificación de datos) se pueden realizar en la nube.

Aunque la Figura 8 ilustra un dispositivo móvil 800, más en general, las innovaciones aquí descritas pueden implementarse con dispositivos que tengan otras capacidades de pantalla y factores de forma de dispositivo, tales como un computadora de escritorio, una pantalla de televisión o dispositivo conectado a una televisión (por ejemplo, una caja de tv por cable o consola de juegos). Los servicios pueden ser proporcionados por la nube a través de proveedores de servicio o a través de otros proveedores de servicios en línea. Además, ya que las tecnologías aquí descritas pueden referirse a retransmisión de audio, una pantalla de dispositivo puede no ser requerida o utilizada (una pantalla puede utilizarse en casos cuando el contenido de audio/video se transmite a un dispositivo de punto final multimedia con capacidades de reproducción de vídeo).

La Figura 9 es un diagrama de un sistema de cómputo ilustrativo, en el que algunas modalidades descritas pueden ser implementadas. El sistema de cómputo 900 no pretende sugerir ninguna limitación en cuanto al alcance del uso o funcionalidad, ya que las innovaciones pueden ser implementadas en diversos sistemas de cómputo de propósito general o propósito general.

Con referencia a la Figura 9, el sistema de cómputo 900 incluye una o más unidades de procesamiento 910, 915 y memoria 920, 925. En la Figura 9, esta configuración básica 930 se incluye

dentro de una línea discontinua. Las unidades de procesamiento 910, 915 ejecutan instrucciones ejecutables por computadora. Una unidad de procesamiento puede ser una unidad de procesamiento central de propósito general (CPU), procesador en un circuito integrado de aplicación específica (ASIC) o cualquier otro tipo de procesador. En un sistema de procesamiento múltiple, múltiples unidades de procesamiento ejecutan instrucciones ejecutables por computadora para aumentar la potencia de procesamiento. Por ejemplo, la Figura 9 muestra una unidad de procesamiento central 910 así como una unidad de procesamiento de gráficos o unidad de co-procesamiento 915. La memoria tangible 920, 925 puede ser memoria volátil (por ejemplo, registros, memoria caché, RAM), memoria no volátil (por ejemplo, ROM, EEPROM, memoria flash, etc.), o alguna combinación de las dos, accesible a través de la unidad(es) de procesamiento. La memoria 920, 925 almacena *software* 980 que implementa una o más innovaciones aquí descritas, en la forma de instrucciones ejecutables por computadora adecuadas para ejecución por la unidad(es) de procesamiento.

Un sistema de cómputo puede también tener funciones adicionales. Por ejemplo, el sistema de cómputo 900 incluye el almacenamiento 940, uno o más dispositivos de entrada 950, uno o más dispositivos de salida 960 y una o más conexiones de comunicación 970. Un mecanismo de interconexión (no mostrado) tal como un conductor común, controlador o red interconecta los componentes del sistema de cómputo 900. Normalmente, el *software*

del sistema operativo (no mostrado) proporciona un entorno operativo para otro *software* que se ejecuta en el sistema de cómputo 900 y coordina las actividades de los componentes del sistema de cómputo 900.

5 El almacenamiento tangible 940 puede ser removible o no removible e incluye discos magnéticos, cintas magnéticas o casetes, CD-ROMs, DVDs, o cualquier otro medio que pueda ser utilizado para almacenar información y que pueda ser accedido en el sistema de cómputo 900. El almacenamiento 940 almacena las instrucciones
10 para el *software* 980 implementando una o más innovaciones aquí descritas.

El dispositivo(s) de entrada 950 puede ser un dispositivo de entrada táctil, tal como un teclado, ratón, lápiz o seguibola, un dispositivo de entrada de voz, un dispositivo de escaneo u otro
15 dispositivo que proporcione una entrada al sistema de cómputo 900. Para la codificación de vídeo, el dispositivo(s) de entrada 950 puede ser una cámara, tarjeta de vídeo, tarjeta sintonizadora de TV o dispositivo similar que acepta la entrada de vídeo en forma analógica o digital, o un CD-ROM o CD-RW que lee muestras de vídeo en el
20 sistema de cómputo 900. El dispositivo(s) de salida 960 puede ser una pantalla, impresora, altavoces, disco compacto grabable u otro dispositivo que proporcione salida desde el sistema de cómputo 900.

La conexión(es) de comunicación 970 permite la comunicación a través de un medio de comunicación a otra entidad de cómputo. El
25 medio de comunicación transmite información tal como instrucciones

ejecutables por computadora, entrada o de salida de audio o vídeo u otros datos en una señal de datos modulada. Una señal de datos modulada es una señal de que tiene una o más de sus características fijadas o cambiadas de tal manera que codifica información en la señal. A modo de ejemplo, y no de limitación, los medios de comunicación pueden usar un portador eléctrico, óptico, RF u otro portador.

Las innovaciones pueden ser descritas en el contexto general de instrucciones ejecutables por computadora, tales como aquellas incluidas en los módulos de programa, que se ejecutan en un sistema de cómputo en un procesador real o virtual objetivo. Generalmente, los módulos de programa incluyen rutinas, programas, bibliotecas, objetos, clases, componentes, estructuras de datos, etc. que realizan determinadas tareas o implementan determinados tipos de datos abstractos. La funcionalidad de los módulos de programa se puede combinar o dividir entre los módulos de programa como se desee en varias modalidades. Las instrucciones ejecutables por computadora para programar los módulos pueden ser ejecutadas dentro de un sistema de cómputo local o distribuido.

Los términos "sistema" y "dispositivo" se utilizan indistintamente en este documento. A menos que el contexto indique claramente lo contrario, ningún término implica ninguna limitación en un tipo de sistema de cómputo o dispositivo de cómputo. En general, un sistema de cómputo o dispositivo de cómputo puede ser local o distribuido y puede incluir cualquier combinación de *hardware* de

propósito especial y/o *hardware* de propósito general con *software* implementando la funcionalidad aquí descrita.

La Figura 10 es un ejemplo de entorno de cómputo de nube que puede utilizarse en conjunción con las tecnologías aquí
5 descritas. El entorno de cómputo de nube 1000 comprende servicios de cómputo de nube 1010. Los servicios de cómputo de nube 1010 pueden comprender varios tipos de recursos de cómputo de nube, tales como servidores, repositorios de almacenamiento de datos, recursos de red, etc. Los servicios de cómputo de nube 1010 pueden
10 estar centralmente ubicados (por ejemplo, proporcionados por un centro de datos de una empresa u organización) o distribuidos (por ejemplo, proporcionados por varios recursos de cómputo localizados en diferentes lugares, tales como centros de datos diferentes y/o localizados en diferentes ciudades o países). Además, el servicio de
15 cómputo de nube 1010 puede implementar la EULT 118 y otras funcionalidades descritas en el presente documento con relación a la actualización de los modelos de clasificador de entendimiento de lenguaje.

Los servicios de cómputo de nube 1010 son utilizados por
20 varios tipos de dispositivos de cómputo (por ejemplo, dispositivos de computación de cliente), tales como los dispositivos de computación 1020, 1022 y 1024. Por ejemplo, los dispositivos de cómputo (por ejemplo, 1020, 1022 y 1024) pueden ser computadoras (por ejemplo, computadoras de escritorio o portátiles), dispositivos móviles (por
25 ejemplo, computadoras de tableta o teléfonos inteligentes) u otros

tipos de dispositivos de cómputo. Por ejemplo, los dispositivos de cómputo (por ejemplo, 1020, 1022 y 1024) pueden utilizar los servicios de cómputo de nube 1010 para realizar operaciones de cómputo (por ejemplo, procesamiento de datos, almacenamiento de datos, generación y edición de definición de agente reactivo, y similares).

En aras de la presentación, la descripción detallada utiliza términos como "determinar" y "usar" para describir las operaciones de cómputo en un sistema de cómputo. Estos términos son abstracciones de alto nivel de operaciones realizadas por una computadora y no deben confundirse con los actos realizados por un ser humano. Las operaciones de computadora reales correspondientes a estos términos varían dependiendo de la implementación.

Aunque las operaciones de algunos de los métodos descritos se describen en un determinado orden secuencial, para una presentación conveniente, debe entenderse que esta forma de descripción abarca la reorganización, a menos que un determinado ordenamiento sea requerido por un lenguaje específico que se establece a continuación. Por ejemplo, las operaciones descritas secuencialmente en algunos casos pueden ser reordenadas o realizadas concurrentemente. Además, en aras de la simplicidad, las figuras adjuntas no pueden mostrar las diversas formas en las que los métodos descritos pueden utilizarse junto con otros métodos.

Cualquiera de los métodos descritos puede ser implementado

como instrucciones ejecutables por computadora o un producto de programa de computadora almacenado en uno o más medios de almacenamiento legibles por computadora y se ejecutan en un dispositivo de cómputo (por ejemplo, cualquier dispositivo de
5 cómputo, incluyendo teléfonos inteligentes u otros dispositivos móviles que incluyen *hardware* computacional). Los medios de almacenamiento legibles por computadora son cualquier medio tangible disponible que se puede acceder dentro de un entorno de cómputo (por ejemplo, uno o más discos de medios ópticos tal como
10 CD o DVD, componentes de memoria volátil (tal como DRAM o SRAM) o componentes de memoria no volátil (tal como memoria flash o unidades de disco duro)). A modo de ejemplo y con referencia a la Figura 9, medios de almacenamiento legibles por computadora incluyen memoria 920 y 925, y almacenamiento 940. El término
15 "medios de almacenamiento legibles por computadora" no incluyen señales y ondas portadoras. Además, el término "medios de almacenamiento legibles por computadora" no incluyen conexiones de comunicación (por ejemplo, 970).

Cualquiera de las instrucciones ejecutables por computadora
20 para implementar las técnicas descritas, así como cualquiera de los datos creados y utilizados durante la ejecución de las modalidades descritas pueden ser almacenados en uno o más medios de almacenamiento legibles por computadora. Las instrucciones ejecutables por computadora pueden ser parte de, por ejemplo, una
25 aplicación de *software* dedicada o una aplicación de *software* que se

accede o descarga a través de un navegador web u otra aplicación de *software* (tal como una aplicación de cómputo remota). Este *software* puede ser ejecutado, por ejemplo, en una sola computadora local (por ejemplo, cualquier computadora comercialmente disponible
5 adecuada) o en un entorno de red (por ejemplo, a través de Internet, una red de área amplia, una red de área local, una red cliente-servidor (tal como una red de cómputo de nube), o cualquier otra red) mediante una o más computadoras de red.

Para mayor claridad, sólo determinados aspectos de las
10 implementaciones basadas en *software* se describen. Otros detalles que son bien conocidos en el arte son omitidos. Por ejemplo, debe entenderse que la tecnología descrita no está limitada a ningún lenguaje o programa de computadora específico. Por ejemplo, la tecnología descrita puede ser implementada por *software* escrito en
15 C++, Java, Perl, JavaScript, Adobe Flash, o cualquier otro lenguaje de programación adecuado. Asimismo, la tecnología descrita no está limitada a ninguna computadora en particular o tipo de *hardware*. Ciertos detalles de computadoras y *hardware* adecuados son bien conocidos y no necesitan ser establecidos en detalle en esta
20 descripción.

Además, ninguna de las modalidades basadas en *software* (que comprende, por ejemplo, instrucciones ejecutables por computadora para hacer que una computadora realice cualquiera de los métodos descritos) se puede cargar, descargar o acceder de forma remota a
25 través de un medio de comunicación adecuado. Tales medios de

comunicación adecuados incluyen, por ejemplo, el Internet, la Gran Red Mundial, una Intranet, aplicaciones de *software*, cable (incluyendo cable de fibra óptica), comunicaciones magnéticas, comunicaciones electromagnéticos (incluyendo comunicaciones RF, microondas e infrarrojas), comunicaciones electrónicas u otros medios de comunicación.

Los métodos, aparatos y sistemas descritos no deben interpretarse como una limitación de ninguna manera. En cambio, la presente descripción se dirige hacia todas las características novedosas y no obvias y aspectos de las varias modalidades descritas, solas y en varias combinaciones y subcombinaciones entre sí. Los métodos, aparatos y sistemas descritos no se limitan a cualquier aspecto o característica específica o combinación de los mismos, ni tampoco las modalidades descritas requieren que cualquier de una o más ventajas específicas estén presentes o los problemas pueden resolverse.

Las tecnologías de cualquier ejemplo pueden ser combinadas con las tecnologías descritas en uno o más de los otros ejemplos. En vista de las muchas posibles modalidades a las cuales los principios de la tecnología descrita pueden aplicarse, debe reconocerse que las modalidades ilustradas son ejemplos de la tecnología descrita y no debe ser tomada como una limitación del alcance de la tecnología descrita. Más bien, el alcance de la tecnología descrita incluye lo que está cubierto por el alcance y el espíritu de las siguientes reivindicaciones.