

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 373 114**

51 Int. Cl.:
G10L 15/22 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 96 Número de solicitud europea: **06708544 .9**
96 Fecha de presentación: **27.02.2006**
97 Número de publicación de la solicitud: **1859438**
97 Fecha de publicación de la solicitud: **28.11.2007**

54 Título: **PROCEDIMIENTO PARA PROPORCIONAR UN SERVICIO DE VOZ INTERACTIVO SOBRE UNA PLATAFORMA ACCESIBLE A UN TERMINAL CLIENTE, SERVICIO DE VOZ, PROGRAMA INFORMÁTICO Y SERVIDOR CORRESPONDIENTES.**

30 Prioridad:
18.03.2005 FR 0502743

45 Fecha de publicación de la mención BOPI:
31.01.2012

45 Fecha de la publicación del folleto de la patente:
31.01.2012

73 Titular/es:
**FRANCE TELECOM
6 PLACE D'ALLERAY
75015 PARIS, FR**

72 Inventor/es:
**PAILLET, Eric;
TEZE, Vincent y
GYSS, Jean François**

74 Agente: **Pérez Barquín, Eliana**

ES 2 373 114 T3

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

DESCRIPCIÓN

Procedimiento para proporcionar un servicio de voz interactivo sobre una plataforma accesible a un terminal cliente, servicio de voz, programa informático y servidor correspondientes

5

1. Campo de la invención

El campo de la invención es el de los servicios de voz interactivos, a los cuales un usuario puede acceder a partir de un teléfono fijo, un teléfono móvil o un ordenador aplicando la tecnología de voz sobre IP.

10

Más concretamente, la invención se refiere a tales servicios basados en el uso de páginas VoiceXML (que es un lenguaje XML (*eXtended Markup Language*) tal como se describe en "*VoiceXML Version 2.0 Specification W3C Recommendation*", 16 de marzo de 2004, <http://www.w3.org/TR/voicexml120/>, normalizado a W3C por el "*World Wide Web Consortium*", y que define interacciones de voz) o de cualquier otro formalismo equivalente, por ejemplo: SALT (*Speech Applications Language Tags*) tal como se describe en "*Speech Applications Language Tags 1.0 Specification*", 15 de julio de 2002, <http://www.saltforum.org/default.asp>, X+V (por XHTML y Voice, siendo X+V un formato de descripción de interacciones multimodales) tal como se describe en "*XHTML+Voice Profile 1.0 Specification W3C Note*", 21 de diciembre de 2001, <http://www.w3.org/TR/xhtml+voice...>

15

20

A título de ejemplo, el centro de llamadas *Service Clients Wanadoo* (marca registrada) proporciona asistencia técnica a los usuarios de *Wanadoo* (marca registrada) mediante un servicio de voz interactivo. Este servicio de voz interactivo recibe llamadas de los usuarios, y a continuación las redirige hacia un servicio o un asesor técnico apropiado. Tal servicio de voz permite, además, proporcionar una información a un usuario, sin intervención de un asesor técnico.

25

2. Soluciones de la técnica anterior

Una técnica bien conocida que permite diseñar y desarrollar servicios de voz interactivos se basa en el uso de páginas VoiceXML.

30

Un documento VoiceXML define interacciones de voz y/o sonoras entre un usuario (desde su teléfono, por ejemplo) y una plataforma de aplicación de un servicio de voz dado que debe ser conducido por el interpretador VoiceXML situado en la plataforma. Una sesión de usuario consiste de este modo en varios diálogos cargados por la plataforma según las respuestas del usuario.

35

Aunque algunos de los formatos de descripción de servicios de voz interactivos están normalizados y han alcanzado una madurez industrial (por ejemplo la versión 2.0 del VoiceXML), pocos proveedores de plataforma integran el conjunto de las recomendaciones de las normas (VoiceXML, SALT, X+V...).

40

En algunos caso, algunos comportamientos difieren de un proveedor de plataforma a otro, o bien porque la norma se ha interpretado de manera diferente, o bien porque la plataforma no está actualizada, es decir, no respeta todas las exigencias de la norma. De este modo, a pesar de estos formatos de descripción normalizados, no existe aún compatibilidad total entre plataformas

45

Asimismo, estos formatos normalizados no ofrecen en general ninguna funcionalidad específica (tratamiento del lenguaje natural, redireccionamiento de las llamadas, etc.), lo cual requiere desarrollos adicionales más o menos complejos en las plataformas

50

Finalmente, en la mayoría de los casos, la sola utilización de estos formatos normalizados, en particular de páginas VoiceXML, es insuficiente ya que los lenguajes que se les asocia, a menudo interpretados, sólo proponen una solución para describir interacciones con recursos de voz (difusión de mensajes guía grabados o sintetizados, reconocimiento de voz) pero no permiten describir la estructura de un diálogo ni las interacciones con el sistema de información. Por mensaje guía, se entiende, según una definición comúnmente admitida, cualquier tipo de guías sonoras difundidas hacia el usuario. Estas guías sonoras pueden generarse por un módulo de síntesis de voz (TTS, *Text To Speech*) o haber sido grabadas previamente en un archivo.

55

60

Por lo tanto, la mayoría de las veces, hay que completar con componentes de software (por ejemplo "*servlets*", o "módulos de extensión de las funciones del servidor de aplicaciones") y páginas dinámicas (por ejemplo JSP (*Java Server Page*), que es una norma que permite desarrollar aplicaciones cuyo contenido es dinámico) en el servidor de aplicaciones (desempeñando la función de sistema de información) para definir completamente el servicio de voz interactivo.

Todos estos puntos complican el diseño, la aplicación y el mantenimiento de los diferentes componentes de los servicios de voz interactivos.

5 Esta complejidad se caracteriza en particular por una fuerte imbricación de los datos relativos a la presentación de las informaciones que se utilizan para el tratamiento y las transiciones entre los diferentes estados del diálogo del servicio de voz.

10 Los documentos D1: DESAI (US 2003/078779 A1), D2: BALL "SPEECH-ENABLED SERVICES USING TELEPORTAL SOFTWARE ND VOICE XML", D3: SIEMENS (EP 0848373 A), D4: UNYSYS CORPORATION (WO 02/091364 A1) y D5: FRANCE TELECOM (US 2005/091958) ilustran el estado de la técnica anterior.

El documento D1 describe la arquitectura de un servidor de respuesta de voz denominado "Anita". También describe un procedimiento de navegación en los diferentes nodos que componen una red a partir de este servidor.

15 El documento D2 es un artículo de prensa relativo a un sistema denominado "TelePortal". Este sistema aplica un marco de desarrollo de los servicios de voz interactivos que integra una separación entre la lógica de negocio y la presentación de la información, del mismo modo que se realiza desde hace tiempo en los sistemas "clásicos" mediante el uso del patrón de diseño MVC.

20 El documento D3 describe un procedimiento y un sistema que permiten convertir documentos HTML en informaciones utilizables por los servicios de voz interactivos con el fin de proponer una distribución dinámica de los contenidos de las páginas web en forma sonora.

25 El documento D4 se refiere a un sistema cliente/servidor que permite mantener una conversación con un usuario mediante una interfaz telefónica. La parte de cliente del sistema integra un navegador capaz de interpretar los documentos VoiceXML. La parte de servidor del sistema incluye un interpretador de flujo de diálogo (DFI) que usa las informaciones proporcionadas por el cliente para generar, dinámicamente, respuestas destinadas al usuario. Los archivos generados incluyen mensajes guía, archivos de gramáticas y otras informaciones.

30 El documento D5 se refiere a la mutualización de la creación de los servicios de voz con el fin de permitir una creación y un mantenimiento fáciles de los servicios. Este documento proporciona un procedimiento de creación de servicios de voz interactivos que es sencillo y que no requiere conocimientos profundos.

35 Los documentos D6 ("*A Portable, Server-Side Dialog Framework for VoiceXML*" de BOB CARPENTER, SASHA CASKEY, KRISHNA DAYANIDHI, CAROLINE DROUIN, ROBERTO PIERACCINI, SPEECHWORKS INTERNATIONAL ET AL., ICSLP 2002: 7TH INTERNATIONAL CONFERENCE ON SPOKEN LANGUAGE PROCESSING, DENVER, COLORADO; (ICSLP), ADELAIDE: CAUSAL PRODUCTIONS, AU, 16 de septiembre de 2002 (16/09/2002), página 2705, X P0070117201 S B N: 978-1-876346-40-9) y D7 ("*Mainstreaming speech-enabled Web applications*" de Juan Huerta, David Lubensky, David Nahamoo, Roberto Pieraccini, T.V. Raman y Charle Wiecha, IBM T.J. Watson Research Center, diciembre de 2004 (14/12/2004), son asimismo documentos, no patentes, que ilustran el estado de la técnica anterior.

45 El documento D6 se refiere a un marco de desarrollo de aplicaciones de voz. El documento D7 se refiere asimismo a un marco de desarrollo de aplicaciones de voz.

50 En particular, el documento D6 describe un archivo de descripción de servicio de voz en la sección "ETUDE DIALOG MANAGER". Este archivo de descripción se presenta en formato UML (fig.2). Un archivo en formato VoiceXML también se crea usando *servlets* genéricos para la gestión estándar del diálogo, acciones y errores (comparar con la sección REPEAT AND BACKUP). La gestión del diálogo se realiza también generando páginas de tipo JSP, y un *servlet* "genérico" de tipo "Aplicación" se utiliza para los componentes de software específicos (recursos de servidor específicos de diálogo). En el documento D7, el 1º y 2º párrafos de la página 6 describen las herramientas necesarias para construir componentes de software específicos a partir de los componentes genéricos.

55 3. Objetivos de la invención

La invención tiene especialmente como objetivo resolver los inconvenientes de la técnica anterior.

60 Más concretamente, un objetivo de la invención es proporcionar una técnica de implementación de servicios de voz interactivos que permita describir estos servicios de manera sencilla, legible y genérica.

Otro objetivo de la invención es proporcionar una técnica que permita describir la estructura de los diálogos y las

interacciones con el sistema de información de manera sencilla y eficaz.

Otro objetivo de la invención es proponer una técnica que permita simplificar el diseño, la aplicación, el mantenimiento y las evoluciones de los servicios de voz interactivos.

5 Otro objetivo adicional de la invención es proponer una técnica que permita reducir notablemente el tiempo de desarrollo de un servicio de voz interactivo.

10 La invención también tiene como objetivo proporcionar una técnica que permita evitar una fuerte imbricación de los datos de presentación que se utilizan para los tratamientos y las transiciones entre los diferentes estados de los diálogos de los servicios de voz.

Otro objetivo adicional de la invención es proponer una técnica que pueda hacerse cargo de manera automática de al menos una parte de los servicios de voz interactivos.

15 **4. Características esenciales de la invención**

Estos objetivos, así como otros que aparecerán más adelante, se alcanzan con la ayuda de un procedimiento que proporciona un servicio de voz interactivo, entre un conjunto de servicios de voz interactivos, en una plataforma accesible para un terminal cliente, tal como se describe con relación a las reivindicaciones 1 a 7.

La invención se refiere asimismo a un servidor según la reivindicación 9 para la aplicación del procedimiento reivindicado, así como a un producto de programa informático según la reivindicación 8.

25 De este modo, se prevé de manera ventajosa aislar la parte genérica de un servicio de voz interactivo en un archivo de descripción.

Por parte genérica se entiende todos los elementos simples y estáticos del servicio, tales como las guías sonoras (sin partes dinámicas), los archivos de gramáticas y las transiciones sencillas o totalmente descritas.

30 En efecto, en la gran mayoría de los servicios de voz hay elementos sencillos y recurrentes que pueden aislarse para describir una parte de los servicios de manera genérica. La invención propone describir esta parte genérica en un archivo de descripción del servicio que puede estar, por ejemplo, en formato XML.

35 La presente invención propone asimismo trasladar los elementos dinámicos o complejos de un servicio de voz interactivo a componentes de software específicos (por ejemplo en *servlets*, JSP...) que son referencias en un archivo de descripción cuando una parte del servicio es dinámica o compleja (en el caso de los mensajes compuestos, por ejemplo).

40 Asimismo, la invención propone automatizar toda o parte de la aplicación de los servicios de voz, facilitar el mantenimiento y las actualizaciones, utilizando un motor de fases que agrupe todos los componentes de software simples que permiten el tratamiento de las partes genéricas descritas en el archivo de descripción.

45 Por motor de fases genérico, se entiende que el motor de fases puede reutilizarse tal cual y sin adaptaciones para diferentes servicios de voz.

Según un aspecto ventajoso de la invención, dicho archivo de descripción asociado a dicho servicio comprende:

- 50
- al menos un elemento estático de dicho servicio interpretable por dicho motor de fases; y/o
 - al menos un referenciamiento hacia dicho al menos un segundo componente de software, específico de dicho servicio, que permita generar al menos un elemento dinámico y/o complejo de dicho servicio

55 De este modo, el archivo de descripción del servicio permite describir la integridad del servicio, o bien de manera directa para los elementos estáticos, o bien recurriendo a componentes de software dedicados para los elementos dinámicos o complejos tales como, por ejemplo, mensajes compuestos, interacciones con un sistema de información...

60 Preferiblemente, dicho segundo componente de software genera dinámicamente al menos una página de diálogo a partir de al menos una norma predeterminada.

De este modo, el segundo componente de software puede aplicar tecnologías (o normas) tales como JSP (*Java*

Server Page), ASP (*Active Server Page*), CGI-BIN (*Common Gateway Interface-Binary*) o PHP (*Personal Home Page*), o cualquier otra tecnología equivalente.

De manera ventajosa, dicho motor de fases comprende medios de:

- 5
- decodificación de dicho archivo de descripción asociado a dicho servicio;
 - suministro de una lista de al menos un mensaje guía;
- 10
- generación de al menos una transición de estado de dicho servicio;
 - carga de al menos una gramática;
 - delegación de al menos un tratamiento a dicho segundo componente de software específico.
- 15
- Preferiblemente, dicho motor de fases comprende:
- 20
- al menos un primer *servlet* genérico que permite la decodificación de dicho archivo de descripción asociado a dicho servicio durante dicha etapa de interpretación;
 - al menos una página dinámica genérica de tipo JSP que genera una página de diálogo de dicho servicio en función de un estado actual de dicho diálogo;
 - al menos un segundo *servlet* genérico que se deriva de dicho primer *servlet* genérico y que permite:
- 25
- proporcionar dicho estado actual a dicha página dinámica genérica;
 - delegar un tratamiento a dicho segundo componente de software específico.
- 30
- De manera ventajosa, dicho motor de fases comprende asimismo un archivo de descripción de características de dicha plataforma.
- En una realización preferida de la invención, dicho procedimiento se aplica en un servidor que realiza igualmente, previa solicitud de dicha plataforma, una etapa de generación de al menos una página de diálogo de dicho servicio a partir de al menos algunos de los siguientes elementos:
- 35
- al menos uno de dichos elementos estáticos;
 - al menos uno de dichos elementos dinámicos y/o complejos;
 - al menos una característica de dicha plataforma extraída de dicho archivo de descripción de características de dicha plataforma.
- 40
- De manera ventajosa, dichas páginas de diálogo son generadas dinámicamente según al menos un formalismo predeterminado de descripción de interacciones.
- 45
- Por formalismo de descripción de interacciones se entiende cualquier tipo de lenguaje que permita definir interacciones (de tipo vocal, multimodales, etc.), por ejemplo VoiceXML, SALT, X+V, XML o cualquier otro formalismo equivalente.
- 50
- La invención se refiere asimismo a un servicio de voz interactivo accesible para un terminal cliente en una plataforma de una red de comunicación, definido mediante un archivo de descripción que comprende:
- 55
- al menos un elemento estático interpretable por un motor de fases, que comprende al menos un primer componente de software, y genérico para al menos dos servicios de voz distintos; y/o
 - al menos un referenciamiento hacia al menos un segundo componente de software, específico de dicho servicio, que permite generar al menos un elemento dinámico y/o complejo de dicho servicio.
- 60
- La invención se refiere asimismo a un producto de programa informático que comprende instrucciones de código de programa para la ejecución de las etapas del procedimiento anteriormente mencionado, cuando dicho programa se ejecuta en y/o por un procesador.

La invención se refiere también a un servidor que comprende medios de generación de al menos una página de diálogo de un servicio de voz interactivo, entre un conjunto de servicios de voz interactivos, accesible en una plataforma para un terminal cliente, y que comprende medios de mando de un motor de fases, que comprenden al menos un primer componente de software, genérico para dicho conjunto de servicios de voz interactivos, y al menos un segundo componente de software específico de dicho servicio, permitiendo dichos medios de mando la interpretación por dicho motor de fases de un archivo de descripción asociado a dicho servicio, de manera que permita dicha generación de dicha página de diálogo.

5. Lista de figuras

Otras características y ventajas de la invención resultarán más evidentes tras analizar la siguiente descripción de una realización preferida, descrita de manera ilustrativa y no limitativa, y los dibujos anexos, en los que:

- la figura 1 es un organigrama que ilustra de manera global un procedimiento para proporcionar un servicio de voz interactivo según una realización preferida de la invención;

- la figura 2 es un esquema simplificado que ilustra el procedimiento para proporcionar el servicio de voz interactivo de la figura 1;

- la figura 3 presenta un esquema que ilustra el funcionamiento del servidor de aplicaciones durante la aplicación del procedimiento para proporcionar el servicio de voz interactivo de la figura 1; y

- la figura 4 presenta la estructura de una realización particular de un servicio de voz según la invención.

6. Descripción de la invención según una realización preferida

6.1 Presentación general de la invención

La invención se basa en la siguiente observación: los elementos de base de los lenguajes de descripción estándar (normalmente los mejor soportados por los diferentes proveedores de plataforma y que constituyen el menor denominador común para la compatibilidad) son también los elementos más usados por los servicios de voz.

De este modo, aunque sea difícil realizar un servicio de voz interactivo de manera genérica, en la gran mayoría de los servicios de voz interactivos existen elementos simples y recurrentes que pueden aislarse para describir una parte del servicio de voz de manera genérica, formando los otros elementos, denominados específicos, una parte específica del servicio de voz.

Por este motivo, la invención propone describir esta parte genérica de un servicio de voz interactivo en un archivo de descripción del servicio (también denominado archivo de descripción de la parte genérica del servicio), que está, por ejemplo, en formato XML.

Los elementos dinámicos o complejos en el servicio de voz se trasladan entonces a componentes de software (por ejemplo en *servlets*) que se referencian en el archivo de descripción del servicio de voz interactivo ya que una parte del servicio es dinámica o compleja (como es el caso, por ejemplo, de los mensajes compuestos).

De una parte del servicio se encarga un componente específico ya que es dinámica, compleja, o se necesitan tratamientos que no podrán efectuarse mediante los lenguajes de descripción estándar (por ejemplo el lenguaje VoiceXML) tales como las interacciones con el sistema de información o módulos de software que efectúan cálculos pesados como vocabularios muy extensos o el lenguaje natural.

De este modo, el servicio está compuesto por dos tipos de componentes de software:

- componentes simples (por ejemplo *servlets*, en adelante designados como *servlets* genéricos) que permiten el tratamiento de las partes genéricas descritas en el archivo de descripción;

- componentes específicos del servicio (por ejemplo *servlets*, en adelante designados como *servlets* específicos).

En consecuencia, la solución de la invención permite, a partir de elementos de descripción del servicio (el archivo de descripción del servicio) en un formato predefinido y por una serie de herramientas adaptadas (por ejemplo *servlets*), automatizar toda o parte de la aplicación de un servicio de voz, y facilitar el mantenimiento y las actualizaciones.

El procedimiento de la invención permite:

- 5 - describir el servicio de voz de manera legible y genérica;
- hacerse cargo de los elementos genéricos de manera automática;
- facilitar el mantenimiento y las evoluciones del servicio de voz;
- 10 - disponer de una arquitectura flexible y evolutiva;
- reducir el tiempo de desarrollo del servicio de voz interactivo, gracias al diseño de una parte genérica común a todos los servicios de voz.

15 6.2 Realización preferida de la invención

En relación con la figura 1 se presenta un organigrama que ilustra de manera global un procedimiento para proporcionar un servicio de voz interactivo según una realización preferida de la invención. Este procedimiento se aplica en un servidor de aplicaciones conectado a una plataforma, mediante una red de comunicación, que permite ejecutar páginas VoiceXML de un diálogo del servicio de voz.

Tras recibir una solicitud 10 utilizando el protocolo http (*HyperText Transfer Protocol* o protocolo de transferencia de hipertexto) solicitando la generación de una página VoiceXML y procedente de la plataforma, un *servlet* genérico del motor de fases realiza una descodificación de las informaciones estáticas 11 comprendidas en un archivo de descripción 12 del servicio de voz interactivo. El archivo de descripción del servicio de voz está escrito en lenguaje XML.

La invención propone por lo tanto el uso de un formalismo de descripción muy sencillo que permite describir los elementos estáticos del servicio y el encadenamiento de las acciones o de las transiciones.

A continuación, si el servicio necesita la aplicación de un tratamiento dinámico o complejo, el archivo de descripción del servicio comprende uno o más referenciamientos a uno o más *servlets* específicos, y un *servlet* genérico delega el tratamiento dinámico o complejo que corresponde a estos *servlets* específicos que realizan una descodificación de las informaciones dinámicas 13.

De este modo, el servidor de aplicaciones se encarga también de los elementos dinámicos, estando el tratamiento correspondiente (acción, generación de guías dinámicas, transición, tratamiento avanzado de los eventos de reconocimiento de voz, etc.) referenciado en el formalismo de descripción constituido por el archivo de descripción del servicio. Se utiliza de este modo el lenguaje Java en un entorno (o arquitectura) de N capas para los tratamientos, donde cada servidor efectúa una tarea especializada (por ejemplo servidor web, servidor de base de datos...). Una resolución de las características dependientes de la plataforma 14 mediante un archivo de descripción de las características de las plataformas 15 se aplica a continuación. El archivo de descripción de las características de las plataformas tiene asimismo formato XML.

Se recuerda que las características de la plataforma son utilizadas por una página JSP (por una página JSP genérica o por una página JSP específica) en el momento de la generación de la página VoiceXML. Las características propias de cada plataforma susceptible de ser accedida se almacenan en un archivo (denominado archivo de descripción de las características de la plataforma). Estas características definen en particular el formato de archivos sonoros aceptados por la plataforma o también el formato de los modelos soportados para las gramáticas de voz.

Todos los datos propios de la plataforma son sustituidos por los elementos genéricos cuya resolución se realiza en el último momento mediante datos de configuración del software.

Finalmente, una página dinámica genérica, que es una página JSP (*Java Server Page*) genérica o, si fuese necesario, una página dinámica específica, que es una página JSP específica, proceden a la generación de la o las páginas VoiceXML solicitadas 16. La o las páginas generadas se envían a continuación 17 a la plataforma.

Los elementos útiles para la plataforma (especialmente las páginas VoiceXML) se generan entonces a voluntad en el servidor de aplicaciones a partir de los elementos estáticos, de los tratamientos dinámicos y de la resolución de las características dependientes de la plataforma.

La generación de estos elementos se realiza mediante un módulo de software que define un tratamiento por defecto y por cuyo comportamiento puede ser sobrecargado (adaptado) por el desarrollador del servicio de voz interactivo mediante mecanismos propios de los lenguajes objeto. Esta sobrecarga del comportamiento permite al desarrollador realizar los tratamientos específicos necesarios para la elaboración de las páginas dinámicas.

5 De esta manera, el desarrollador del servicio dispone de un medio que le permite concentrarse solamente en las partes dinámicas del servicio, siendo el resto gestionado por el archivo de descripción genérica y el motor de fases.

10 Por lo tanto, los tratamientos se mutualizan de manera muy sencilla, lo cual permite reducir los costes de aplicación y de mantenimiento de los servicios de voz.

Asimismo, el formalismo de descripción y las herramientas utilizadas por la invención podrán permitir tomar en cuenta los aspectos de multimodalidad de un servicio.

15 Tal y como se ilustra en la figura 2, que presenta un esquema simplificado del procedimiento para proporcionar el servicio de voz interactivo de la figura 1, la solución propuesta por la invención comprende dos tipos de componentes:

20 - un primer tipo de componente 201 de definición del servicio de voz interactivo mediante un archivo de descripción del servicio 12 en formato XML, de un conjunto de *servlets* específicos 23 y de una página JSP específica 26 y

25 - un segundo tipo de componente 202 de ejecución del servicio de voz, que aplica un conjunto de componentes de software (agrupados bajo la denominación de motor de fases 20) que comprende *servlets* genéricos, una página JSP genérica y el archivo de descripción de las características de las plataformas, que le permite hacerse cargo de manera automática de los elementos anteriormente mencionados que definen el servicio.

En lo sucesivo se describe el archivo 12 de descripción del servicio anteriormente mencionado.

30 La idea es referenciar en este archivo de descripción de servicio solamente los elementos simples y fijos como las guías sonoras (si no contienen partes dinámicas), los archivos de gramáticas o también los nombres de los *servlets* específicos encargados de realizar las transiciones.

35 Puesto que un elemento o una transición pueden describirse totalmente de manera única y sin ambigüedad, se pueden referenciar en el archivo de descripción.

En todos los otros casos, el archivo de descripción del servicio recurre a *servlets* específicos que se encargan de la producción de los elementos dinámicos y/o condicionales (por ejemplo guías sonoras condicionadas por datos de contexto, guías sonoras cuyo contenido es dinámico, transiciones condicionadas por datos de contexto, etc.).

40 Se obtiene de este modo una descripción del servicio de voz interactivo legible (por lo tanto adaptado al mantenimiento o a las evoluciones de manera sencilla) y fácil de general de manera automatizada.

En el anexo 1 se presenta un ejemplo del contenido del archivo de descripción del servicio. Este anexo forma parte integrante de la presente descripción.

45 Este ejemplo describe dos fases que son «Bienvenida» y «Menú». Estas dos fases describen interacciones de tipo vocal (materializadas en el archivo mediante «<MEDIA type="VOICE" id=" 1 ">»). De este modo se considera una extensión del formalismo para poder describir interacciones relativas a otros medios y, más generalmente, para las aplicaciones multimodales.

50 En la fase de «Bienvenida», el sistema comienza difundiendo una guía sonora (cuyo identificador es 236) no permitiendo al usuario interrumpir la guía antes de su finalización (lo que se traduce en el archivo por «bargein="false"»). El texto asociado a la guía en el archivo de descripción del servicio se proporciona en caso de que el sistema no encuentre el identificador en un archivo de indexación de las guías.

55 El archivo de indexación de las guías asocia cada identificador a un archivo de sonido a difundir y a un texto a sintetizar si el archivo es inaccesible. En este caso se difundirá el texto «Bienvenido al servicio de voz» en síntesis de voz.

60 A continuación, el sistema efectúa una ramificación hacia un *servlet* específico denominado «VxmlServletSuiteBienvenue» que se encarga de generar la página VoiceXML siguiente.

En la página «Menú», el sistema efectúa las siguientes tareas:

- llamada de un *servlet* *específico* denominado «VxmlServletGuideMenu» encargado de la difusión de uno o más mensajes guía (guías sonoras);

5

- carga de una gramática «gram_estrella» (de tipo DTMF, *Dual Tone Multi Frequency*);

- carga de una gramática «recol.srg» (de tipo reconocimiento de voz);

10

- elección de la transición a activar.

Es posible añadir atributos para las guías y las gramáticas además de los directamente gestionados por el formalismo. Para esto, basta con una actualización de una DTD (*Document Type Definition*, definición del tipo de documento) ya que todos los atributos no gestionados directamente por el formalismo (por ejemplo, los identificadores, el tipo de gramática...) se retomarán tal cual en la página VoiceXML generada.

15

El paso de un estado a otro se define por el elemento «<TRANSITION>» del estado actual y puede llevarse a cabo de varias maneras.

20

El paso puede llevarse a cabo directamente hacia otro estado del servicio; en este caso el *servlet* genérico no abstracto descrito posteriormente es llamado por la página JSP genérica cuando se evalúa el elemento «<SUBMIT>».

25

Asimismo puede definirse el siguiente estado mediante un *servlet* que será llamado por la página JSP cuando se evalúe el elemento «<SUBMIT>».

También puede definirse el siguiente estado mediante condiciones que pueden ser de varios tipos:

30

- evento «NO INPUT» (sin acción del usuario);

- evento «NO MATCH» (acción no tratada por el sistema);

- condición sobre el estado anterior;

35

- condición gestionada por un *servlet* en cuyo caso el tratamiento del estado se delega a este *servlet*;

- condición sobre un evento con nombre (por ejemplo "connection disconnect" o "error"): su prioridad de evaluación depende del formalismo usado en la salida (VoiceXML u otro);

40

- condición por defecto: sólo se evalúa si todas las otras condiciones son falsas;

- condición sobre evento de reconocimiento de tipo vocal o DTMF.

45

Cada condición (la condición gestionada por un *servlet*) contiene a su vez un elemento «<TRANSITION>» que se cumple si se toma en cuenta la condición.

50

Puede observarse que el número de delegación para los *servlets* está empíricamente limitado a 2 (es decir que un *servlet* que ya se ha llamado directamente desde otro *servlet* no podrá a su vez delegar el tratamiento en otro *servlet*). Aunque esto permite cubrir la gran mayoría de los casos, puede imponer restricciones sobre la escritura del archivo de descripción del servicio.

A continuación se describen los *servlets* específicos y la página JSP específica mencionados anteriormente.

55

Se recuerda que en algunos casos puede ser necesario recurrir a componentes de software para tratar los datos dinámicos, para recurrir a tratamientos externos al sistema, para acceder a una base de datos, etc.

Estos componentes de software se implementan en forma de *servlets* específicos y de la página JSP específica.

60

Los componentes de software pueden encargarse, por ejemplo, de generar guías sonoras dinámicas (cuyo contenido depende, por ejemplo, de la hora actual) o también de determinar la siguiente página VoiceXML a difundir en función de elementos tales como el resultado de un análisis semántico, el valor de un parámetro contenido en una base de datos, etc.

De este modo, la parte específica del servicio se compone del conjunto de los *servlets* específicos y de la página JSP específica. Los *servlets* específicos son referenciados en el archivo de descripción del servicio en la etapa correspondiente al tratamiento específico.

5 Los *servlets* específicos se derivan de los *servlets* genéricos de manera que pueden tratar, sin reescribir el código, toda la parte genérica de la etapa correspondiente. Solo la parte de tratamiento del resultado de la gramática necesita la escritura de código; esto se realiza en el procedimiento *getNextState*.

10 Se recuerda que el procedimiento *getNextState()* permite determinar el próximo estado en la descripción del servicio, es decir, la próxima transición.

Se obtiene de este modo gracias al uso de los conceptos de herencia y de sobrecarga de los lenguajes orientados a objetos (Java), una simplificación de los tratamientos específicos.

15 Los estados que cargan una gramática deberán usar un *servlet* específico para la transición. Este *servlet* específico permite recuperar el resultado de la gramática, hacer el análisis de la misma y hacer la transición hacia el estado correspondiente.

20 En la mayoría de los casos, los *servlets* específicos se utilizan:

- para gestionar la lista de los mensajes guía a ejecutar; en este caso sólo hay que redefinir el procedimiento que crea esta lista;

25 - para gestionar una transición; en este caso sólo hay que redefinir el procedimiento que devuelve el nombre del estado siguiente.

En el archivo de descripción del servicio es posible indicar algunos atributos con nombres de constantes (por ejemplo: *timeout="T0"*) para las guías de audio y para las gramáticas. En este caso, un *servlet* específico determinado (o una de las clases de la jerarquía) debe definir una variable denominada T0.

30 Durante la llamada a *getGuidesFromDescripciónState()*, el *servlet* genérico (que forma parte de la jerarquía del *servlet* específico determinado) busca la constante denominada T0 por el mecanismo de reflexión o de introspección (mecanismo de la programación orientada a objetos que permite descubrir los procedimientos y las variables accesibles en un objeto particular, sin conocimientos a priori de este objeto). Si se encuentra esta constante, se utiliza su valor, sino se deja el valor T0 como valor del atributo.

Hay que resaltar que el procedimiento *getGuidesFromDescripciónState()* permite buscar las guías sonoras a difundir a partir de la descripción del estado actual en el archivo de descripción del servicio.

40 A continuación se describe el motor de fases anteriormente citado.

Se utilizan páginas VoiceXML rudimentarias, que comprenden:

45 - la difusión de guía(s);

- la carga de gramáticas DTMF o de reconocimiento de voz;

- la espera de una acción de un usuario;

50 - la emisión de una solicitud hacia el servidor de aplicaciones para obtener la siguiente página VoiceXML.

El uso de páginas rudimentarias se impone ya que se quiere utilizar tratamientos particulares entre dos interacciones del usuario. Este es el caso, por ejemplo, en lenguaje natural cuando debe hacerse un análisis semántico de la frase devuelta por el motor de reconocimiento con el fin de decidir cual es la siguiente fase a activar.

55

Estas páginas rudimentarias permiten asimismo reducir el número de etiquetas (o marcadores) VXML utilizadas y, por lo tanto, limitar los riesgos sobre la compatibilidad entre las diferentes plataformas.

60 La parte genérica puede interpretarse por un conjunto de *servlets* genéricos reutilizables para todos los servicios. Estos *servlets* genéricos son capaces de interpretar el archivo de descripción del servicio cuando no referencian

servlets específicos. En el caso en que un *servlet* específico es referenciado, los *servlets* genéricos son capaces de delegar el tratamiento a este *servlet* específico.

5 El motor de fases se basa principalmente en dos *servlets* genéricos y en la página JSP genérica anteriormente mencionados.

10 El motor de fases comprende un primer *servlet* genérico que permite la lectura y la descodificación del archivo de descripción del servicio en el procedimiento de inicialización. La descripción del servicio se almacena en un objeto miembro de manera independiente del formato del archivo de descripción. En este caso, el acceso es más rápido durante la ejecución propia del servicio que una búsqueda de las informaciones en el árbol DOM (*Document Object Model*) correspondiente al archivo de descripción del servicio.

15 Este comprende asimismo un segundo *servlet* genérico que se deriva del primero y que es capaz de leer el contenido de un estado y de preparar la página JSP genérica o bien delegar el tratamiento a un eventual *servlet* específico. Los elementos buscados en el estado son:

- la lista de los mensajes guía a ejecutar o el nombre del *servlet* específico si la lista de mensajes guía es sometida a condiciones complejas (como informaciones de perfil del llamante);

20 - la lista de las gramáticas a cargar;

- la lista de tratamientos de tipo "*no inputs*", que puede ser una lista de mensajes guía o una página a cargar (es decir un *servlet* o una página JSP o también una página VoiceXML);

25 - la página a cargar en un evento de tipo "*no match*";

- la página a cargar en un evento denominado («*bloc "catch"*»);

30 - la página a cargar según un resultado de la gramática, o la página siguiente si no hay gramática que cargar;

- el nombre del estado siguiente si puede describirse directamente en el archivo de descripción del servicio o el *servlet* específico capaz de encontrar el estado siguiente.

35 Esto comprende asimismo la página JSP genérica que permite generar una página VoiceXML utilizando las informaciones encontradas en el estado actual. El paso al estado siguiente se hace por la evaluación de una baliza «*<SUBMIT>*» cuya fuente corresponde a la página siguiente indicada por el *servlet* genérico no abstracto durante la llamada de la página JSP genérica.

40 El motor de fases comprende finalmente el archivo de descripción de las características de las plataformas mencionado anteriormente.

Con relación a la figura 3 se presenta el funcionamiento del servidor de aplicaciones durante la implementación del procedimiento para proporcionar el servicio de voz interactivo de la figura 1.

45 De este modo, al recibir una solicitud procedente de la plataforma y que comprende, por ejemplo, informaciones sobre un estado actual, un estado siguiente (si se conoce), una causa de la llamada que inicia la ejecución del servicio de voz interactivo, un resultado de reconocimiento de voz, un nombre de evento, el *servlet* genérico 31 del motor de fases descodifica el archivo de descripción de servicio 12.

50 Se generan elementos dinámicos y/o complejos del servicio por *servlets* específicos 23 referenciados en el archivo de descripción de servicio 12. Algunos de los *servlets* específicos pueden usar un analizador semántico 33 en el marco, por ejemplo, de la aplicación del lenguaje natural.

55 La técnica de la herencia 34 puede aplicarse al nivel de los *servlets* específicos 23.

La página JSP genérica 35, o la JSP específica 26, genera una página VoiceXML 38 del diálogo a partir:

- de información sobre un estado actual y sobre un estado siguiente (si se conoce);

60 - de una lista de las guías;

- de las gramáticas;

- de los DTMF;
- de las palabras de mando;
- 5 - de una lista de los eventos a gestionar;
- de una lista de las guías y/o página siguiente sobre inactividad;
- 10 - de una página siguiente sobre rechazo («no-match»);
- de al menos una característica de la plataforma, extraída del archivo de descripción de características de plataforma 15.

15 La página VoiceXML así gestionada se envía a continuación a la plataforma.

El servlet genérico 31, la JSP genérica 35 y el archivo de descripción de características de plataforma 15 forman parte integrante del motor de fases. Estos elementos no cambian de un servicio a otro.

20 El archivo 12 de descripción del servicio, los *servlets* específicos 23 y la o las JSP específicas 26 son propios de cada servicio y definen los elementos complejos o dinámicos del servicio.

25 La figura 4 presenta finalmente la estructura de un servicio de voz 52 según la invención, que comprende una memoria 522, y una unidad de tratamiento 521 equipada con un microprocesador μ P, que está controlado por un programa informático (o aplicación) 523. La unidad de tratamiento 521 recibe como entrada una solicitud 51 procedente de una plataforma. El microprocesador μ P trata esta solicitud, según las instrucciones del programa 523, para generar páginas VoiceXML, que son transmitidas hacia la plataforma 53.

30 En resumen, la invención propone un procedimiento para proporcionar un servicio de voz a partir de un archivo de descripción (por ejemplo de tipo XML) que comprende elementos genéricos que describen una parte del servicio de voz, y componentes de software específicos, que actúan conjuntamente con un motor de fases. Este procedimiento presenta numerosas ventajas, tales como descorrelacionar eficazmente las partes estáticas y las partes dinámicas de un servicio de voz, de manera que facilita el desarrollo y el mantenimiento del mismo. En efecto, en el caso de una evolución, una actualización o también una corrección del servicio de voz, el desarrollador del servicio sólo trata

35 las partes dinámicas. La invención se aplica especialmente a todos los tipos de servicios de voz: los servicios en lenguaje natural, en la detección de palabras clave (por ejemplo según la técnica «*multi-word spotting*» o «detección de palabras clave»), en palabras envueltas o aisladas, pero también en los servicios que utilizan las interacciones con las teclas del teléfono (por ejemplo DTMF).

40 **Anexo 1**

```

<STATE name="Bienvenida">
  <MEDIA type="VOICE" id="1">
    <GUIDE bargein="false" id="236">Bienvenido al servicio de voz</GUIDE>
    <TRANSITION type="SERVLET_MANAGED">VxmlServletSuiteBienvenue</TRANSITION>
  </MEDIA>
</STATE>
<STATE name="Menú">
  <MEDIA type="VOICE" id=" 1">
    <GUIDE id="SERVLET_MANAGED">VxmlServletGuideMenu</GUIDE>
    <GRAMMAR type="DTMF" 1d="1">gram_estrella</GRAMMAR>
    <GRAMMAR type="VOICE" id="2">reco1.srg</GRAMMAR>
    <TRANSITION type="CONDICIONAL">
      <TRANSCOND type="NO-INPUT" id=" 1 ">
        <TRANSITION type="STATE">Menú del sitio</TRANSITION>
      </TRANSCOND>
      <TRANSCOND type="DTMF" value="" id="2">
        <TRANSITION type="SERVLET_MANAGED">VxmlServletMenuDTMF</TRANSITION>
      </TRANSCOND>
      <TRANSCOND type="EVENT" value="connection.disconnect" id="3">
        <TRANSITION type="STATE">Desconexión</TRANSITION>
      </TRANSCOND>
  </MEDIA>
</STATE>

```

```
<TRANSCOND type="DEFAULT" id="4">  
  <TRANSITION type="SERVLET_MANAGED">VxmlServletMenuReco</TRANSITION>  
</TRANSCOND>  
<TRANSITION>  
5 </MEDIA>  
</STATE>
```

REIVINDICACIONES

1. Procedimiento de puesta a disposición de un servicio de voz interactivo en formato VoiceXML, de entre un conjunto de servicios de voz interactivos, en una plataforma accesible para un terminal cliente, constituyendo dicha plataforma un elemento intermedio entre dicho terminal y un servidor; caracterizado porque dicho servicio de voz interactivo comprende un archivo de descripción (12) que describe dicho servicio y al menos un componente de software específico a dicho servicio; cuando dicho terminal cliente accede a dicho servicio de voz interactivo, el procedimiento comprende una etapa de interpretación de dicho archivo de descripción (12) y de dicho al menos un componente de software específico a dicho servicio, comprendiendo además dicho procedimiento:
- una etapa para generar una página de diálogo de dicho servicio, en formato VoiceXML, según un estado actual de dicho diálogo y de dicho archivo de descripción (12);
 - una etapa para delegar una operación de procesamiento a realizar por dicho servicio de voz en dicho componente de software específico según dicho estado actual de dicho diálogo y de dicho archivo de descripción (12);
- y porque dicho servidor accede, durante dicha etapa de generación y dicha etapa de delegación, a un archivo de descripción (15) que describe características de dicha plataforma, con el fin de que dicho servicio de voz interactivo se adapte a dicha plataforma.
2. Procedimiento de suministro según la reivindicación 1, caracterizado porque dicho archivo de descripción (12) asociado a dicho servicio comprende:
- al menos un elemento estático (11) de dicho servicio que puede interpretarse por un motor de fases ; y/o
 - al menos un referenciamiento hacia al menos un segundo componente de software específico a dicho servicio, que permite generar al menos un elemento dinámico y/o complejo (13) de dicho servicio.
3. Procedimiento de suministro según la reivindicación 2, caracterizado porque dicho componente de software genera dinámicamente al menos una página de diálogo a partir de al menos una norma predeterminada.
4. Procedimiento de suministro según la reivindicación 1, caracterizado porque se implementa mediante un motor de fases que es genérico para dicho conjunto de servicios de voz interactivos.
5. Procedimiento de suministro según una cualquiera de las reivindicaciones 3 y 4, caracterizado porque dicho motor de fases (20) comprende medios para:
- decodificar dicho archivo de descripción (12) asociado a dicho servicio;
 - suministrar una lista de al menos un mensaje guía;
 - generar al menos una transición de estado de dicho servicio;
 - cargar al menos una gramática;
 - delegar al menos una operación de procesamiento en dicho al menos un componente de software específico.
6. Procedimiento de suministro según la reivindicación 1, caracterizado porque dicho procedimiento se implementa en un servidor que también lleva a cabo, previa solicitud (10) de dicha plataforma, una etapa de generación de al menos una página de diálogo de dicho servicio a partir de al menos algunos de dichos elementos siguientes:
- al menos uno de dichos elementos estáticos (11);
 - al menos uno de dichos elementos dinámicos y/o complejos (13);
 - al menos una característica de dicha plataforma, extraída de dicho archivo de descripción (15) que describe características de dicha plataforma.
7. Procedimiento de suministro según la reivindicación 1, caracterizado porque dichas páginas de diálogo se generan dinámicamente según al menos un formalismo predeterminado de descripción de interacciones.
8. Producto de programa informático (523) que comprende instrucciones de código de programa para la ejecución

de las etapas del procedimiento para proporcionar un servicio de voz interactivo según una cualquiera de las reivindicaciones 1 a 7 cuando dicho programa se ejecuta en y/o por un procesador.

- 5 9. Servidor que comprende medios de generación de al menos una página de diálogo de un servicio de voz interactivo en formato VoiceXML, de entre un conjunto de servicios de voz interactivos, accesible en una plataforma para un terminal cliente, constituyendo dicha plataforma un elemento intermedio entre dicho terminal y dicho servidor; caracterizado porque dicho servidor comprende además medios de mando que permiten interpretar un archivo de descripción (12) que describe dicho servicio y al menos un componente de software específico a dicho servicio, con el fin de permitir dicha generación de dicha página de diálogo, comprendiendo además dicho servidor:
- 10 - medios para generar una página de diálogo de dicho servicio, en formato VoiceXML, según un estado actual de dicho diálogo y de dicho archivo de descripción (12);
- 15 - medios para delegar una operación de procesamiento a realizar por dicho servicio de voz en dicho componente de software específico según dicho estado actual de dicho diálogo y de dicho archivo de descripción (12);
- y porque dicho servidor comprende medios para acceder, durante dicha generación y dicha delegación, a un archivo de descripción (15) que describe características de dicha plataforma, con el fin de que dicho servicio de voz interactivo se adapte a dicha plataforma.

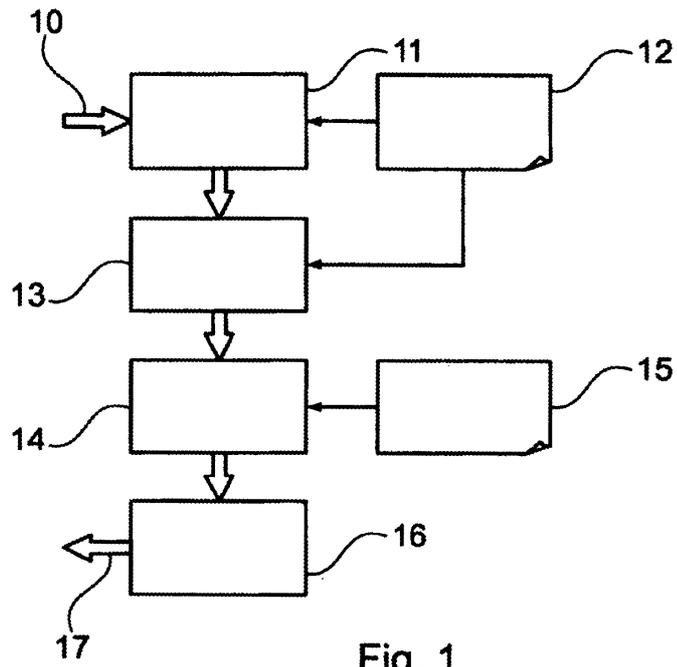


Fig. 1

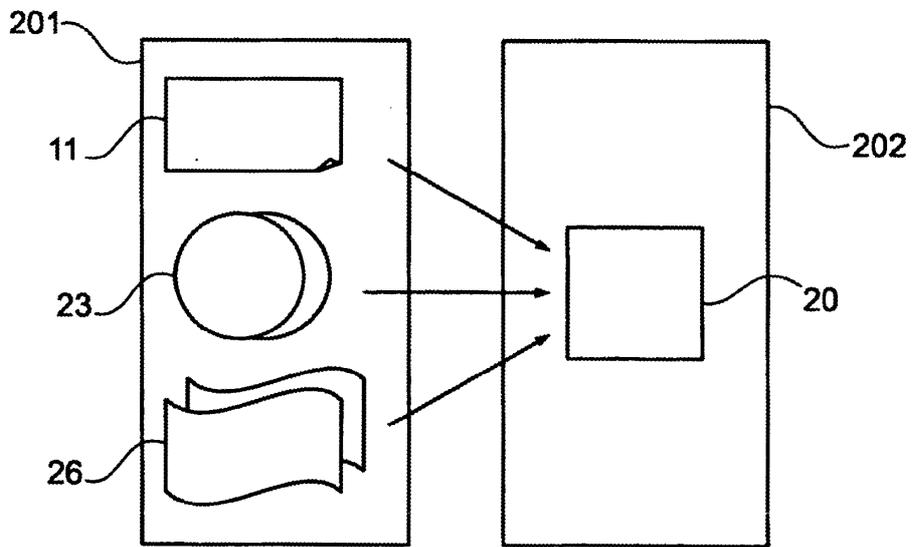


Fig. 2

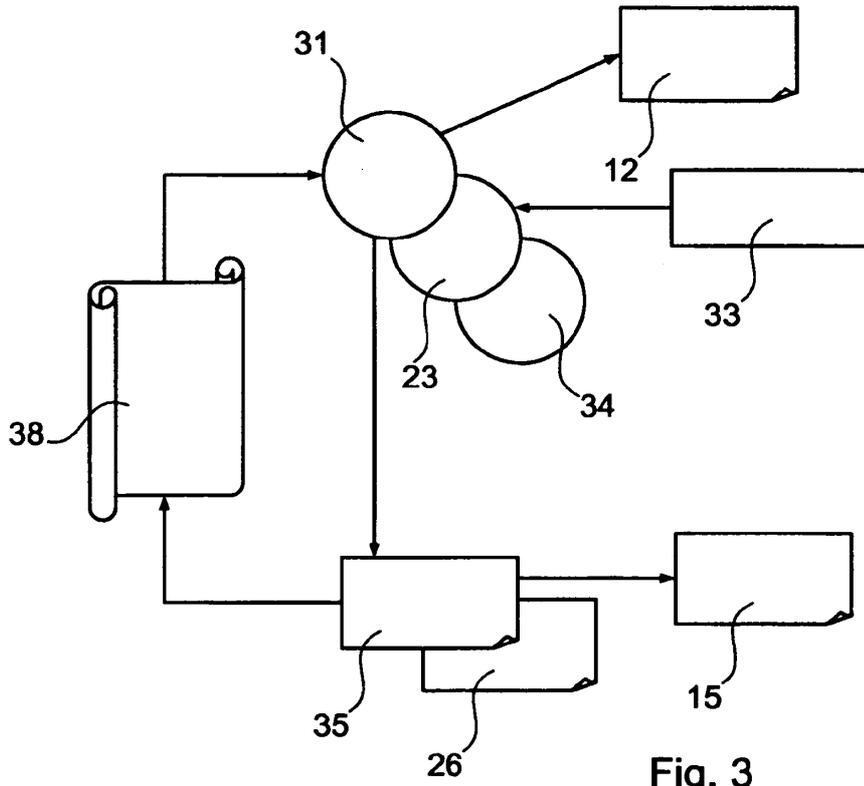


Fig. 3

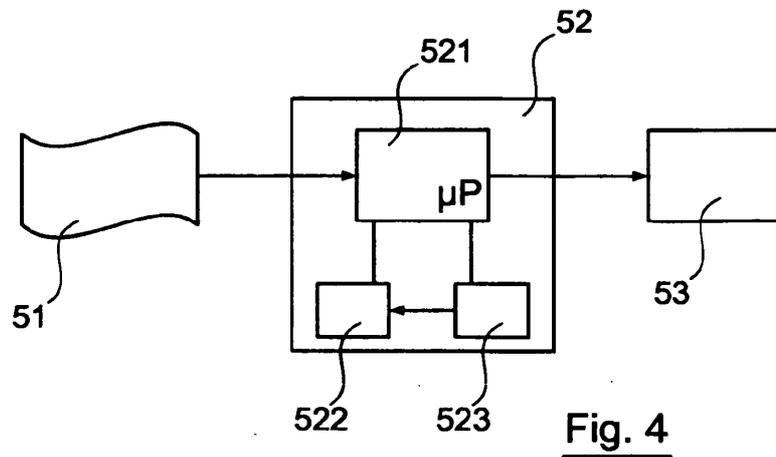


Fig. 4