



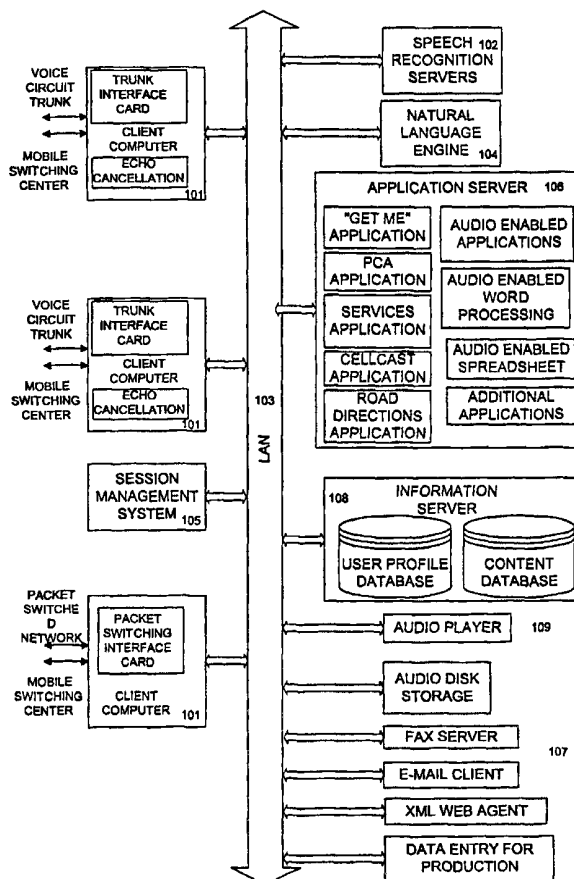
INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification ⁷ : G06F	A2	(11) International Publication Number: WO 00/67091 (43) International Publication Date: 9 November 2000 (09.11.00)
(21) International Application Number: PCT/IL00/00246 (22) International Filing Date: 30 April 2000 (30.04.00) (30) Priority Data: 60/131,491 29 April 1999 (29.04.99) US (71) Applicant (for all designated States except US): SPINTRONICS LTD. [IL/IL]; Kaufman Street 2, 68012 Tel-Aviv (IL). (71)(72) Applicants and Inventors: TE-ENI, Benjamin [IL/IL]; Ir-Shemesh Street 53A, 69068 Tel-Aviv (IL). ISRAELI, Gil [IL/IL]; Hagefen Street 38, 99797 Karme-Yosef (IL). (74) Common Representative: ISRAELI, Gil; Hagefen Street 38, 99797 Karme-Yosef (IL).		(81) Designated States: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG). Published Without international search report and to be republished upon receipt of that report.

(54) Title: INFORMATION RETRIEVAL SYSTEM

(57) Abstract

A platform for accessing and interacting with information according to users' individual interests over a cellular network while using an improved speech interface with natural understanding capabilities, employing a variety of software applications.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INFORMATION RETRIEVAL SYSTEM

1. FIELD OF THE INVENTION

The present invention relates generally to computer based information and
5 commerce systems and wireless communications systems and more particularly,
to a platform for organizing accessing and interacting with information according
to users' individual interests over the cellular network. The present invention thus
transcends the boundaries of radio, Internet, and more traditional information
formats.

10 As described below, the present invention provides a platform that enables access
to information, advertising, and special features and services via the cellular
network in an audio format. The platform is represented by the concept of a
Mobile Agent (a virtual window through which users may access all available
information and services) and of a Virtual Companion (a representation of the
15 system's audio interface by a user-selected persona). The Companion, whose
voice and personality are selected by the user, serves as a guide to the many
services and information retrieval options available to the user via the Mobile
Agent.

Primary operative characteristics of the preferred embodiment of the invention
20 described below are as follows:

For information distribution and retrieval, users access a virtual private hard disk,
which is the user's personal information storage space and database, via a cellular
switching center;

The invention makes use of the cellular phone as a remote input/output (I/O)
25 device, and employs a speech-user-interface enabling speech activation and data
retrieval in a convenient audio format; and

The system enables the information dissemination center to add hyperlinks relevant to radio and TV broadcasts; the hyperlinks are available via the cellular phone, without modifying the actual radio or TV broadcast. Thus users of the system can obtain information regarding media broadcasts via the cellular phone.

- 5 The Mobile Agent functions as a value-added service (VAS) to cellular telephone systems. Value-added services help effect a virtual transmutation of the cellular phone, extending the range of functions from telephony to voice-mail and faxing capabilities, and ultimately far beyond, as the cellular phone morphs into a mobile personal computer (or personal digital assistant). The present invention advances
10 this transmutation by means of a unique speech user-interface (SUI) utilizing speech recognition technologies and innovative audio services, thereby allowing the cellular phone a range of functions restricted heretofore to personal computers.

2. BACKGROUND OF THE INVENTION

- Wireless communications technologies, together with the "Information
15 Superhighway" and World Wide Web, are revolutionizing the ways in which people access personal data, news, and information. The tools now available to consumers are already rapidly changing the ways in which news and information are disseminated, and are altering people's personal habits of data storage and retrieval. Even as people are becoming increasingly mobile, the importance of
20 maintaining a viable link with personal data, together with access to multiple forms of news and information, is becoming increasingly critical. In order to meet the growing demand for personalized, portable, and flexible news and information services, a hybridization of Internet technologies and wireless network capabilities is required
- 25 The trend towards unification and integration of data and communications functions is a direct function of this requirement. The combination of Internet-based services with wireless communications technologies represents the surest means of meeting the demand for mobile, consistent, and reliable data storage, processing, and retrieval.

Communication and data management must be made efficient in terms of both time and money. The rapid proliferation of available data can make retrieval of only the desired items a time-consuming process. Similarly, the sophisticated gadgetry used to perform such functions is a drain on cash and resources.

- 5 While the complicated devices now on the market serve some of these purposes, users are subject to the mechanical limitations of any given device. And even the most useful and clever of such devices may quickly become obsolete.

For news and information, consumers still rely heavily on traditional news media such as print media, radio, and television, although an increasing number of
10 consumers are making use of interactive Web-based technologies for information retrieval purposes.

Traditional news media such as print media, radio, and television have specific disadvantages that can be best mitigated by complementary use of newer technologies. Traditional news media are neither interactive nor personalized to
15 accommodate user preferences, and so lend themselves to inefficiency; time, money, and resources are wasted. For instance, each of these traditional media is restrictive to users, subjecting users to inconvenient physical and temporal restrictions, where all users receive the same broadcast, at the same time. This requires the media consumer's presence either at a particular time or at a specified
20 location – or both (e.g., near the television at 8 o'clock), and enables the consumer an experience limited to the boundaries of the media broadcast planned for all the audience, with virtually no possibility of personalization or interaction with the broadcast. Compared with computerized technologies, traditional media update information slowly and infrequently.

- 25 Not all of these deficiencies have been remedied by the advent of the computerized technologies. For example, one of the Information Age's ironies is that the "virtual" technologies that dominate the news media landscape demand no less than total "actual" presence. Although information dwells in cyberspace, users are expected to be physically present – keying-in commands, pointing and

clicking, visually monitoring a display unit – in order to access information. Consumers are seeking alternatives to the immobile, display-intensive media (such as Internet broadcast networks) that are presently available.

The present invention addresses these problems, rendering the ironies of
5 information retrieval in cyberspace obsolete by allowing users to free themselves from their PC's and workstations without compromising their unfettered access to news and information in cyberspace and the advantages in the availability of existing mass media. An innovative speech user interface (SUI) is most congenial to a fast-paced, mobile lifestyle. Users need not be present at any particular
10 location or at any particular time in order to access news and information items.

Other problems with the new Information Age news media include delays for downloading files, and slow, inefficient, or clumsy user interfaces.

Another problem with existing technologies is the lack of integration of information dissemination and retrieval processes with consumers' lifestyles and
15 habits. Personal data and communications are becoming an increasingly critical aspect of personal and business practices. In order to keep appointments, maintain contacts, and send and receive messages in various forms, consumers must support a range of information databases. The process of managing disparate forms of data and information is technically and logistically cumbersome. A
20 successful information management system will combine personal information such as contact data and appointment information with news stories and entertainment features under the aegis of a single platform.

In addition, existing technologies are not sufficiently adaptable to consumers' behavior, habits, and styles. Although some applications allow users to set
25 preferences, in many cases this is an unnecessary step that could be circumvented by using "smart" technologies that predict user habits.

The present invention addresses the above problems and issues, integrating the capabilities of wireless communications and Internet-based information dissemination systems in a manner not formerly employed.

5 A goal of this invention is to integrate multiple forms of access to electronic data, information, and advertising into a single, easily negotiable format – a Mobile Agent. Users are provided with a wide range of services and information retrieval formats via the cellular network using an audio interface.

10 Another goal of the present invention is to make all processes and services simple and easy to use. The speech-user-interface is coupled with audio-formatted data files, and then endowed with specific user-defined characteristics. Having selected a specific persona for the interface, the user interacts with the resultant Companion, which serves as a guide to all services and forms of information available via the Mobile Agent.

15 Another goal of the present invention is to provide extensive personalization and customization options. The Companion adapts to each user's habits and preferences, and recognizes each user's voice and speech patterns. Users select a persona for their Companion, set preferences for all system functions and services, and create a self-profile that helps determine which advertising items are played to the user and in what manner and conditions. The system adapts to
20 patterns of typical use.

Another goal of this invention is to provide users with customization options such that the user may determine which categories of news and information items the user would like to receive and not receive. Using the present system, users are able to set and change preferences either over the wireless network using a
25 cellular handset, or alternatively, over the World Wide Web using a PC.

Another goal of this invention is to provide information access in an audio format. The user is not required to refer to a visual display or GUI in order to access the desired information during the information retrieval process.

Another goal of the present invention is to provide access to news and information such that access to information is completely "hands-free." The user is not required to "key-in" input commands, touch or manipulate an input device in order to retrieve information. The user interacts with information such as

5 broadcast radio and TV by selecting from a list of links respective to a specific media program. Thus the user is able to receive additional information regarding such programs or items on the program and purchase the information or commercial items related to the program. All using voice or WAP menu commands.

10 Another goal of the present invention is to assure each subscriber timely information upon request without delays for downloading files. Upon hearing a song the subscriber likes, he can immediately download the song to his mobile phone using speech or WAP commands to facilitate the download.

Another goal of the present invention is to present information and news items to

15 users in formats of varying lengths so that information is disseminated at a rate that suits each user's preferences.

Another goal of the present invention is to provide quick, specific answers to straightforward questions. (e.g., "what is the song played on Capital Radio right now"? and : where can I buy the product being advertised right now on the radio?)

20 Another goal of the present invention is to provide road directions to mobile phone users, with integrated up-to-the-minute data such as traffic and road-condition information with road directions.

Another goal of the present invention is to augment the utility of information services by combining personal assistant and organizer functions with news and

25 information dissemination. The user is able to keep track of media bookmarks. The system stores a time stamp upon receiving the user's 'bookmark' command. The bookmark then enables the user to return to that time frame on a specific

radio or TV station, thereby relating to a playlist of news items, music items, or advertisements broadcasted at that time on the selected station.

- Another goal of the present invention is to actively facilitate commerce and financial transactions via the cellular network. Another goal of the present invention is to provide remote access to applications, files, and data, together with extensive storage space, via the Virtual PC. System based application settings, files, and data are synchronized with information on the user's PC such that copies of user files may be obtained via the cellular system in audio format or via the user's PC.
- 10 Another goal of this invention is to conserve time, money, and the resources involved in communications and data access and retrieval, while providing news, information, and advertisements in a dynamic, efficient, entertaining format.

3. SUMMARY OF PRESENT INVENTION

- The present invention describes an information system for organizing accessing and interacting with information and media according to users' individual interests over the cellular network comprising: a speech recognition engine for converting speech received from the subscribers cellular telephone handset to to a plurality of commands for operations to be performed according to the speech from the subscribers cellular telephone handset; a Natural Language Engine for compiling the required speech interface and sending it to the recognition engine; a session management system for directing commands from recognition engine to the appropriate application and database typically via an application server, enabling entering and retrieving data from the respective application database; a Profile database for storing, updating, and retrieving personal data regarding preferences of the respective subscriber regarding content, and speech interface, and a Content database for storing data as required by the subscriber enabling such data to be retrieved as desired by the subscriber.

and an array of value added service applications designed for cellular phones.

It should be noticed that both the session management and the application server may be implemented on the same server computer and commercial application server software platform such as Oracle Application Server, or BEA Weblogic

According to further features in the described preferred embodiment, the
5 application server integrates:

- A Get Me application and database for ordering goods and services including information items, media programs, and other commercial products according to the respective command; The system functions as an “agent” empowered to buy products according to the user’s commands. These operations may be executed
10 according to a profile of pre-defined user-preferences or system options, and in online, real-time transactions from radio stations/TV, etc. Via the cellular network users may access travel agent services, order a pizza using only a cellular phone, buy a product that is being advertised on the radio or TV, or buy in-depth an Interview that was broadcasted on the radio or TV.
- 15 A News dissemination application and database containing news information to be supplied to the subscribers according to the respective command; The information made available to the users is retrieved from the radio TV and news papers and It is aggregated and tagged with properties and relevant links by a team of writers, editors, and media experts. Where required the information or
20 links are then recorded in an audio format by voice and sound specialists. Human-voice audio clips are then stored on the server’s database.

Subscriber profile data stored in the database. This profile data represents user preferences, which are indications of categories of information items to which the user would like access. The preference controls function as a filter for excluding
25 those items inconsistent with the user profile data from the information items played over the cellular network. Profile data is used to select advertisements appropriate to each user according to the chosen items. Users may set or change these settings using the cellular handset or alternatively, over the Internet using a PC.

The system also provides a system for answering user-initiated queries via the cellular network. To activate these general information services, users may ask questions on a wide range of topics using speech commands. Using advanced query applications, human speech is converted to a query with standardized
5 formulation by a natural language server. Once converted, queries are processed and answered in a convenient audio format. The user can request that a copy of the response will be sent to an email address or fax number.

a Virtual PC mode of remote access to the subscribers' Virtual PC information and applications. The Virtual PC consists of the user's private virtual hard disk
10 (data storage space designated to each user,) together with access to audio enabled applications such as word processing. Users may create directories and sub-directories for storing, organizing, and in which to receive information items. Information (data, files, and application settings) on the Virtual PC is synchronized with information on the user's own PC, such that subscribers are
15 granted remote access to the information stored on their PC together with access audio-enabled files via the cellular network.

A road directions transportation application. Based on natural language server capabilities, the system uses segmented audio clips of road directions stored at a local database. In response to users' requests, individual clips of audio directions
20 are linked together and played for the user via the cellular network. Road directions are customized to suit each user's mode of transportation, listening preferences, chosen route, and personal profile.

A General Information database containing reference information, tools for calculation, games, and other general information, to be supplied to subscribers
25 according to the respective command.

The invention also provides an integrated audio interface called Virtual Companion, a conceptual representation of the user-interface in the form of a persona with specific attributes and characteristics. The interface is based on speech commands as the primary mode of data input and audio files as the

primary mode of data output. Communication with the user is made possible as the system learns user typical responses and requests. Using both text-to-speech and speech-to-text engines, together with advanced voice-recognition technologies, the Virtual Companion enables users to make requests using natural speech commands, and allows users to hear information in a convenient, easy-to-use format.

An audio Mobile Agent system stores and updates a database of information items. The information items are classified according to a precise categorization system whereby each item is tagged according to multiple properties such as source and time of broadcast. This tagging thus allows the system to match articles and advertisements to users, to cater to user preferences most effectively to effect a real-time retrieval of recent articles, advertisements, music items etc. Tagging information items may also include links to related information items or commercial offerings enabling users to access additional information relevant to specific news items or broadcasts, and conduct pertinent commerce transactions.

According to further features in the described preferred embodiment, the system further comprises interfaces for external devices, including an audio player, an audio disk storage, a fax server, and an e-mail client, all controlled by the session management system.

It should be noticed that the natural language engine may be used independently to enhance and cache an existing speech interface used for any application. The Natural Language Engine pre - loads the requested voice interface data or document from a document site or database, performs a series of rule-based scripts and processing functions on the document, and then provides the subscriber with the speech interface to the site requested enhanced by processing functions, all controlled by the session management system.

4. BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram illustrating a preferred system architecture in accordance with the present invention;

Fig. 2 is a block diagram illustrating the system structure;

- 5 Fig. 3 is a diagram illustrating the processes that generate Cellcast summary reports

Fig. 4 is a block diagram illustrating the operations of the Virtual PC

Fig. 5 is a flowchart illustrating the operations of the Adaptive Interface

- 10 Fig. 6 is an expanded illustration of FIG. 5, illustrating an example of system adaptive processes

Fig. 7 is a tree-diagram illustrating an example of user-created directories

Figs. 8a and 8b are flowcharts illustrating the operations performed in the course of a sample phone call

15

Fig. 9 is a block diagram illustrating one form of Natural Language Engine that may be provided to improve speed recognition and understanding performance; and

- 20 Fig. 10 is a diagram illustrating how the Natural Language Engine interacts with the input source types.

5. DESCRIPTION OF THE PREFERRED EMBODIMENT

5.1 System Architecture

- 25 The representative embodiment illustrated in Fig. 1 is a system in which speech recognition engine is comprised of two tiers of computers. The first tier includes computers with telephony cards including hardware and software for interfacing with a telephony network (e.g. Dialogic DTI 300SC-E1 telephony card). The

second tier of computers includes speech recognition servers (102), running the recognition software itself. The computers in this first tier interfacing directly with the telephone network on one end, and employing the speech recognition servers on the other end are referred to hereinafter as "Client Computers" (101). It should
5 be noticed that this architecture is not mandatory for the implementation of the application and the speech recognition server may also reside on a DSP located on the telephony card itself or on the same computer running the telephony interface (e.g. Aculab Prosody cards).

The Client Computers are connected via communication lines to a telephony
10 switching center. The connection to the cellular operator may be via a dedicated connection such as an E1 trunk that carries multiple circuit calls. It is also possible to use a packet switching network for connection to the switching center in order to facilitate access to packet switching enabled cellular phones. Mobile subscribers dial a dedicated access number (e.g., *99) and their call is routed via
15 said trunk to one of the client computers. Each client computer utilizes a trunk termination card that provides the interface between the lines and the client computer's resources. The client computer may also incorporate an echo cancellation card to improve speech quality, enable better recognition rates and cut speech segments in accordance with speaker pauses.

20 All client computers connect to a plurality of servers via a Local Area Network (103) by utilizing open and proprietary protocols such as TCP/IP and Microsoft networking. Digitized speech segments are transferred to a speech recognition server (102), which applies various well-known algorithms for conversion into textual information.

25 The Natural Language Engine (104) retrieves from the database data to be accessed by users using a speech interface, implements several processing modules in order to construct the dialog file, pre-compiles the Recognition Server 102 Grammar and sends the files to the Recognition Server 102. as more particularly described below with respect to Figs. 9 and 10.

The session management system (105) directs commands from recognition engine to the appropriate application and database typically via an application server .

The session management system (105) tracks all ongoing sessions and provides the context for each user service request. The session management system (105)

5 also interfaces with a mechanism for billing each cellular subscriber for its access to system resources. The session management system has access to all applications and resources available on the LAN and provides means for linking all system functions and resources so as to supply a seamless flow of information to and from the user while minimizing unnecessary delays. Among the available

10 resources is an application server (106), which runs dedicated software modules, each designed to handle a subset of the embodiments in accordance with the invention. Further resources include interfaces, generally designated (107), to external devices or computer systems, including a messaging gateway (fax, e-mail, SMS, WAP etc.), a web gateway to facilitate Internet surfing capabilities,

15 and various other value-added gateways. A plurality of database structures (108) is maintained in order to provide access to subscriber profiles and content records.

As an example for a typical session, where a user asks for a Pizza to be sent to his home address from a take-away shop, we may analyze the following session:

(a) Subscriber dials the access number (e.g. *99) and is routed via one of

20 the available trunks to a client computer 101; (b) the client computer software identifies an incoming call and answers it; (c) the session management system 105 handles the call from now on by loading the user's profile from the profile database 108 according to its caller ID data; (d) the natural language engine 104 compiles the dialog interface

25 and send it to the recognition server 102. (e) the corresponding prompt is played by the audio player 109; (e) the user says "get me a Pizza"; (f) said speech segment is sent to the speech recognition server 102, which in turn responds with set of system commands, such as PURCHASE ITEM, where ITEM=PIZZA according to the grammar

30 definitions received from the natural language engine 104; (h) upon

occurrence of a PURCHASE command, the session manager 105 transfers the command to the application server 106, while appending user identity code; (i) the application server searches the user profile database 108 for a definition of the purchase items, such as a Pizza, and in case no such item exists, a response is played back to the user which can select to connect to a pizza shop via the internet through the XML web agent 107. If a valid purchase item is located, further environment variables are looked into. These may include the user's current home address (or current location if pizza is to be sent to his current location which may be provided by the mobile switching center or by the user in speech form). Other implied information elements such as preferred pizza shop, preferred type of pizza etc. may be completed through the user's profile database or through alternative environment variables.

15

5.2 Description of Mobile Agent characteristics:

5.2.1 Natural Language Engine

20 In order to achieve improved speech recognition and understanding performance, the system utilizes a Natural Language Engine (NLE) as illustrated in figure 9.

The Natural Language Engine interacts with an Automatic Speech Recognition (ASR) engine of choice, a Voice Extensible Mark Up Language (VXML) Engine, or any other Voice Markup Language (VML) recognition engine of choice such as the Motorola VoXML engine (all engines also referred to hereinafter as "Recognition Engine"). The Natural Language Engine retrieves the data or document that is to be accessed with a speech interface, i.e. any Voice Markup Language (VML), XML, or other data residing in a database, then implements several processing modules, pre-compiles the required Recognition Engine Grammar and saves it in a grammar cache storage unit. When the data or

25
30

document is accessed by a user via the recognition engine the NLE recognition engine is fed with the preprocessed speech recognition data and grammar files and with the recorded audio prompt files, which make up the system audible response in the Dialogue ("prompts"). The grammar and prompt files are preprocessed in
5 accordance with predetermined rules and artificial intelligence algorithms.

NLE STRUCTURE

The Natural Language Engine, as illustrated in figure 9, consists of several modules:

Document Processing Engine

10 The **Document Processing Engine** (902) is responsible for receiving the voice Dialogue data from the application server or external XML agents, enhancing the Dialogue characteristics and transforming the data to a Dialogue data format understandable by the speech recognition server such as the recognition engine Application Programming Interface (API), VXML or other Dialogue definition
15 language.

The Document Processing Engine includes several sub-modules:

DATA RETRIEVAL UNIT

The **Data Retrieval Unit** (905) Receives data requests from the **NLE Manager** and executes them, i.e. load the required document or information element from
20 the database or from external agents such as the XML agent. The **Data Retrieval Unit** also receives data requests from subsequent layers of the Document Processing Engine (i.e. the **Parsing Unit** and the **Dialogue Enhancement Unit**) to fetch external documents and executes them as well.

PARSING UNIT

25 The **Parsing Unit** (906) receives new documents from the Data Retrieval Unit and converts each document to its own respective internal representation. It consists of several pluggable document parsers; each parser supports a specific

input format (for example VXML, or JSGF), and converts it to an internal XML tree structure representing components of the source document. The parser separates complex document structures to its basic components, to be later processed (and cached) by different modules. Such structure separation occurs, for
5 example, in large VXML files, where an inline grammar is transformed to an external JSGF grammar file.

DIALOGUE ENHANCEMENT UNIT

The **Dialogue Enhancement Unit** (907) processes the XML components supplied by the **Parsing Unit**. It consists of several **enhancing modules**, each adapted to
10 enhance a specific format or data source, such as grammar, Dialogue, prompt, etc. The XML components are processed by separate **enhancing modules** according to their type. Each enhancing module runs a different set of algorithms or scripts designed specifically to its known component type. The enhancing module functions are detailed later in the functional description section. An output format
15 type specifier is attached to each processed XML component by its respective enhancing module, to indicate the possible output formats that can be used by the **Format Wrapper**. Note that some XML components might be created, modified, and destroyed in the enhancing process.

FORMAT WRAPPER

20 After each component has passed through the enhancement process, it is passed to the **Format Wrapper** (908), which transforms it to an output format compatible with the speech recognition server technology. The Format Wrapper consists of several **Formatting Modules**, each supporting a specific type of output format. Some of the key formatting modules are: a **Grammar Compiler**, which compiles
25 grammars to the proprietary ASR API / format in order to speed up recognition time; a **VXML compiler**, which generates standard VXML documents for use of a VXML engine; compilers for other proprietary voice dialogue standards, such as the proprietary Motorola VoXML format; a **Prompt Compiler**, which prepares voice prompts using a Text-To-Speech Engine or from a recorded source etc.

Note that this process might include several stages. For example: a grammar file could be converted from JSGF format to a proprietary grammar format which is compatible with the grammar compiler used by the ASR, and then converted from this proprietary format to a compiled form. Note also that the formatting process
5 might invoke external components (for example an external grammar compiler for the ASR platform).

Caching Server

Each compiled component which is ready to be sent to the Speech Recognition Server is stored in the **NLE Caching Server** (903), with its respective creation
10 timestamp and expiry timestamp. The **NLE Caching Server** acts in a similar way to any caching/proxy server, serving cached documents to its client and managing document invalidation/refresh.

The overall NLE response time and performance is improved by the storage of precompiled grammar files. The NLE Cache receives the compiled Dialogue files
15 from the Dialogue file compiler and stores the complex grammars and prompts thereby relieving a VXML engine from slow and lengthy grammar compilation and speeding up recognition performance in the implementation platform.

The **Caching Server** acts as the one and only interface to the Speech Recognition Server. For each document requested from the ASR platform, the **Caching Server**
20 checks if a valid copy of the document exists in the document cache database. If so – the document is fetched and transferred to the ASR platform (via TCP/IP or any other API required by the ASR). If no valid document is found (i.e. “cache miss”) – the **Caching Server** triggers the **NLE Manager**, which in turn loads the required data from its respective source, runs it through the **Document**
25 **Processing Engine**, and stores a valid copy of the document in the **Caching Server**. The **Caching Server** then immediately fetches this document to the requesting party.

NLE Manager

The **NLE manager** (904) invokes the enhancement process or any part thereof periodically or subsequent a content update in a database or a change in one of the VXML or other documents the system is required to enhance access to. Each

- 5 access to the **NLE Caching Server** is also registered in the **NLE Manager**, which uses this information to predict document requests by the ASR, and prefetch these documents before they're actually requested in order to speed up document fetching time.

- 10 The system can receive a variety of input source types and process them in a variety of forms, as depicted in figure 10. The output format could be VXML, VoXML, or any other proprietary format required by the Speech Recognition platform. For example, given an example VXML source as follows:

```

15 <?xml version="1.0"?>
    <vxml version="1.0">
        <menu>
            <prompt>Which album would you like to buy?
            <enumerate/> </prompt>
            <choice next="madonna.vxml" >
20         Ray of Light by Madonna </choice>
            <choice next="petshopboys.vxml">
Standing on the Shoulder of Giants by Oasis </choice>
            <choice next="richie.vxml">
Back To Front by Lionel Richie </choice>
25     </menu>
    </vxml>

```

The system could transform the source to the following VXML target, with the complex pre-compiled grammar supplied as an external file "**album.gram**".

```

<?xml version="1.0"?>
<vxml version="1.0">
  <form>
5      <field name="album_name">
        <prompt>Which album would you like to
        buy? </prompt>
        <grammar src="album.gram"
        type="application/x-jsgf"/>
10      <block>
          <submit next="buy_album.asp"/>
        </block>
      </field>
    </form>
15 </vxml>

```

The Dialogue Enhancement Unit

The NLE receives the actual Option list available to the user at any given stage of the Dialogue, adds possible grammar options so a variety of voice commands and syntaxes can be recognized with improved performance.

- 20 The actual additions to grammar vary according to the Dialogue content and application purpose. Examples for grammar enhancements include:

a. Global and state specific or data specific prefix/suffix

The NLE may add suitable or generic prefixes such as: "Hmm...", "I want...", "I will go for the...", "I'll take the...", "I would like the..." etc.

- 25 Similarly the NLE may add suffixes such as "...please", "...thank you".

b. Alternative description of items on option list:

NLE can add several descriptions that would be acceptable as a reference to an item on the option list. For example in a “Get me” interactive media application which enables the user to get shows or songs played on the radio when the prompt
5 reads out the recent tracks song list from which the user is expected to select a song that was played on the radio, where the basic grammar would be the name of the song and the performer (e.g. “Rain by Madonna”), the NLE adds grammar that enables the user to say “the song played by Madonna”, “the first song on the list”.

c. Multiple grammar fields in one sentence input:

10 The NLE can add grammar making it possible for the user to answer several questions in the Dialogue at once. For example instead of requiring the user to say name of the song first (e.g. “Rain by Madonna”), then the media (“mp file” or “CD”) and then the form of delivery (“email” or “FedEx”) the user can say: “MP file, CD, email” all at once.

15 *d. Sentence understanding*

NLE can add grammar to enable complete sentence syntax understanding, where in the GETME a song example the user can say: “I would like to buy Madonna’s song as an MP file, please send it to my email” instead of answering any directed questions or using specific grammar and syntax.

20 *e. Ambiguity detection*

The NLE detects ambiguous items in option lists, system-level commands and document-level commands (for example: when the user says ‘help’ he could select the command “help” or the song item “help”). The system applies state specific rules preferring state specific grammar and adds a clarification sub-form
25 in the current document for handling ambiguous command, so that the user is asked to clarify her selection (system: “did you mean the command ‘help’ or the song ‘Help by The Beatles’”).

f. Enhanced navigation commands

The NLE detects the Dialogue structure and adds helpful navigation commands, especially (but not limited to) the “back” command. Additional navigation commands include next, skip, and generic ‘goto’ commands.

5 *g. Foreign Accent adaptation*

In order to improve recognition where users speak with a foreign accent when commanding the speech interface in language which is not their mother tongue (e.g a Frenchman saying “ I would like to buy Stevie Wonder’s CD”) , the NLE retrieves data in certain language and processes the speech commands to phonetic
10 representation in the speech commands original language, transforms the phonetic representation to foreign language accent, and rebuilds the dictionary and grammar of the speech commands.

h. Prompt enhancement

The NLE converts TTS (text to speech) prompts to natural (recorded) prompts on
15 the fly, without requiring the VML content editor to change all its VML documents. The system keeps a local database of matching recorded prompts, which replace the TTS prompts are inserted as part of the TTS prompts.

The NLE can precompile TTS prompts and store them in a local database as recorded (TTS) prompts.

20 In addition, the system automatically adds pauses and intonation tags (as defined in Java Speech Markup Language, for example) according to predefined rules; thereby making the system prompts more fluent and natural sounding.

In addition, the system can automatically convert between different types of audio formats, grammar formats and other external file formats according to the file
25 types required by the target platform.

The Natural Language Engine can be used in order to enable improved speech access to any content site, and can be installed as part of the user site, VML

Engine, and as an independent server the service provider utilizes in the same manner an Internet Service Provider uses a proxy server for internet access.

5.2.2 *Speech Interface, WAP and Hybrid interfaces*

The user typically connects to the Mobile Agent system by dialing a specified
5 toll-free phone number. Upon connection with the server, the user's identity is established. Certain resources can be accessed only after user identity is established using voice user name, voice recognition technologies and a secret password. Once the user's identity has been verified, the audio transmission commences. In the preferred embodiment, the user is guided through the personal
10 audio transmission with a series of voice prompts; or the user may initiate communications using spoken commands. In alternate embodiments, the user communicates using the keypad of the cellular handset. As an added feature, in some systems the speech-based user interface is coordinated to the specific functions of the user-handset, such that information may be displayed on the
15 handset's screen and such that users may key-in commands when appropriate, or scroll through available options. These options are used in conjunction with spoken commands., using WAP (Wireless Application Protocol).

5.2.3 *Virtual Companion*

In the preferred embodiment, users select a persona from a list of predefined
20 settings for the speech-based user interface. The selection of the persona determines the following: (1) voice; (2) tone; (3) content and style of the system voice-prompts.

For instance, a wise old man, a sexy woman, or a cowboy may be selected as a persona, each with appropriate greetings, accents, and comments. Personas may
25 also be constructed using clips of recordings of celebrity voices or famous characters. Pre-recorded audio-clips corresponding to the user-selected persona setting constitute the platform's primary mode of data output; this Virtual Companion serves as a guide to all system functions and services.

Persona settings may include motivational messages, sayings, jokes, or slogans as part of the persona format. Other persona options include daily Bible readings or readings from other religious texts, or other messages which change daily. These messages are interspersed throughout the audio transmission at regular intervals,
5 or during idle time.

Persona settings may be configured to active mode. In active mode, the system will ring the user to give messages, or will prompt the user after a pre-defined period of inactivity (e.g., if the subscriber has not accessed the system for three days, he receives a call from his "Companion" asking how he is, what is new,
10 etc.) In the active mode, the Companion may have mood changes from day to day, or may have "needs" which the subscriber is asked to provide for. For example, the sexy woman Companion may "need" to be "complimented" as a condition for providing help to the user. In this way, the user forms a virtual relationship with the Companion, which serves to bond the user to the system, as
15 the system serves social functions above and beyond the information management functions and services that the system performs.

Similarly, the Companion reacts to changes in user habits from day to day. The system is sensitive to such changes by virtue of the adaptive interface. The Companion asks about recognized changes in user habits and responds to them.
20 For example, the Companion may note that the user has not called Mary in one week, whereas he had previously placed daily calls to Mary. The system may inquire, "What is going on with Mary? Would you like to call her?"

In the preferred embodiment, users configure persona settings using the cellular handset or via the Internet using a PC or workstation.

25 In another embodiment, the persona settings are user-defined. Users may determine which words are uttered for any given command. For example, users may choose a special greeting for themselves. Messages may also be included as part of a user-defined persona. Corporations may use this feature to audio transmission corporate messages to its users.

5.2.4 Adaptive Interface

In the preferred embodiment shown in Fig. 5, the system adapts to user habits by analyzing statistics based on users' listening habits and habits of use. As shown in Fig. 5, the system identifies patterns of regular use according to predefined
5 formulae, and measures the occurrence of identified patterns; if the rate of recurrence of the pattern is above a predefined level, the system "adapts" by setting the identified pattern as the default mode of operations.

In the preferred embodiment, users evaluate the relevance of an information item, the duration of the item, and are able to evaluate the relative level of depth or
10 understanding which the user brings to an item. For example, a user interested in law requests stories related to family law. The evaluative features of the Mobile Agent enable the user to determine at which level of complexity he would like to receive stories, (beginner, experienced user, advanced, professional) the initial level definition results in a set of default interest rating user properties. Users also
15 provide feedback to specific stories. The feedback is registered in the user's database and the matching software adds properties to the initial level definition provided by the user, making the appropriate changes.

An embodiment of the adaptive processes of this invention includes first, recording the subscriber usage data in terms of information item requests from a
20 subscriber. The process compiles the usage data to give a complete usage picture for a given subscriber during a given period of time. Finally, the process compiles the usage data, compares the result with the subscriber's original profile and then adjusts the subscriber profile accordingly. This process assumes that records are tracked by day and by category structure; information item retrievals are tracked
25 by subscriber and by time period; and profile category structure priorities are tracked for each subscriber.

The retrieval system of this invention also includes a process for adjusting subscriber profiles through the introduction of peripheral category structures into their profile from time to time. Subscribers initially create their own profiles by

selecting their relevant areas of interest. As time passes, they refine their profiles directly through relevance feedback, and usage feedback by ordering full-text records from delivered briefs. From each method the subscriber indicates what they like or dislike of what they have received. However, no such feedback is
5 available about records subscribers did not receive. To avoid distorted profiles, the automatic retrieval system of this invention provides a process for occasionally introducing, at defined times or randomly, peripheral category structures into a subscriber's profile to determine if the subscriber's interests are expanding into these peripheral areas. In this way, subscribers get to sample, on a
10 limited basis, emerging fields and have their profiles adjusted automatically.

User profile adjustment includes ranking a subscriber's category structures in order of the number of information items retrieved to determine a usage rank. The usage rank is compared with the original rank of the category structure. Next, a new profile rank is determined for each of N category structures by assigning
15 various rates to the different category structures. Finally, the new ranking for each category structure is determined by summing different ranks for that category structure to determine its new priority value. Rules can be applied to avoid wild swings in profile contents by for example, preventing a category structure from moving more than one place in priority for a given usage.

20 The system is capable of "learning" preferred orders of operation as well as specific data items. That is, the system learns to perform operations in the order the user prefers, and also learns to interpret the user's use of specific terms. For example, if a user repeatedly requests a summary report of CNN upon accessing the Mobile Agent, the system will suggest to the user that a summary report of
25 CNN be set as the default mode of operations. In this way, the user will hear the report without having to request it specifically. An example of this operation is illustrated in Fig. 6.

For another example, if a user repeatedly requests “Ben and Jerry’s” upon asking the system to “get me” ice cream, the system will learn to procure Ben and Jerry’s without being asked to specifically.

- 5 If the recurrence of a pattern is high, but not above a pre-defined threshold for adaptations, the system prompts the user to register his preference with the system. For instance, if the user repeatedly requests Ben and Jerry’s, but also requests other brands of ice cream from time to time, the system will ask the user whether or not Ben and Jerry’s should be set as the default definition for “ice cream.”
- 10 The system also adapts to user speech patterns. If the user tends to use commands other than the system-defined commands to perform certain functions, the system will learn the commands. Advanced forms of speech-recognition technologies may be used to accommodate colloquial language, irregular speech patterns, and background noise.
- 15 Users may choose whether or not they wish to be consulted as to whether or not the system adapts to a given pattern.

- In the preferred embodiment, system adaptations are also enabled by means of analysis of subscribers’ evaluations of articles, advertising, and services. Users are periodically asked to evaluate information items, advertising, or services, either in
- 20 the form of a questionnaire or series of questionnaires, or in the form of a short prompt after a given segment of play. For example, after hearing a Premium report from Cellcast, a user is asked to provide rank the material as “very relevant, somewhat relevant, or not at all relevant.”, and to rank the form of the material as “too long”, “good”, or “too short”. Answers to such system-initiated
- 25 queries are stored together with user profile information, and are used to enable the system to adapt to user preferences.

Embodiments of the retrieval system of this invention can also include enhanced customization and duplicate elimination based upon information item properties.

A subscriber can define certain properties that the subscriber always wants to see, or always wants to discard. Through attribute selections, different subscribers can receive different treatment of the same news event. In other cases, a subscriber may want to see all treatments of a particular event or related to a particular party
5 from all sources (e.g., where a public relations department may want to track all treatments of a particular client by the press).

Additional system adaptive capabilities include but are not limited to the following: users may set preferences for the style of interface by selecting a Virtual Companion; users determine the content of all forms of Cellcast reports,
10 including categories of information items to be received or not received; and advertising information is selected based on a user profile.

5.2.5 *User Profile*

In the preferred embodiment, users complete a profile report upon registration to the Mobile Agent. Data is stored in the system database, (8a, Fig. 1) relating to
15 users' personal details, such as name, contact information, age, gender, place of residence, occupation, income, and areas of interest. Users are also requested to complete a questionnaire, which relates to fields of user interest and affiliation. This information is used to select appropriate advertising information, and supports adaptive and customization options.

20 The step of supplying profile information includes providing the system with a user selection of category records (e.g., selecting Cellcast channels, per Fig. 3) The selected category records are weighted to indicate not only priority among categories, but also degrees of preference. The user may select default weights. If default weights are selected, the system assigns successive decreasing values for
25 the weights based on his or her preferences. Alternatively, the user may enter weights for the various categories. The final weight determination is then made, performed using a pre-defined formula.

In the preferred embodiment, the user profile includes information relating to all system functions and settings. User profile information includes but is not limited

to user preferences relating to Mobile Agent functions and services, Cellcast channels and preferences, settings for the Virtual Companion, and user favorites.

User profile data also includes statistics related to subscribers' listening habits, such as the average duration of calls and the operations performed during each
5 call. Records of all user activities are included, as well.

Users may change their profile information using a cellular phone or via the Internet using a PC.

5.2.6 Advertising

In the preferred embodiment, advertising information is disseminated to users via
10 the Mobile Agent. Advertising items may be played one item at a time or with several separate segments linked together, played sequentially.

Audio advertising information is ranked for relevancy for best advertising segmentation and performance, according to several parameters:

- (1) Topic (i.e., relevancy to user interest topics)
- 15 (2) Target user profile information
- (3) Relevancy to specific locations
- (4) Campaign rules (e.g., Pizza adverts to be played only on Fridays, after 6 PM)
- (4) Duration (i.e., time needed to play each item).
- (5) Mobile Agent functions to which items are best suited.
- 20 When to play which advertising items is determined using a calculus of the above factors.

In the preferred embodiment, users have a choice of service plans. The different plans are classified as Basic, Premium, and Super-premium. Super-premium services contain no advertising. Premium users may make use of an advertising
25 filter in order to request that certain topics of advertising be included or not

included in the Mobile Agent services. Basic service users receive advertising at a higher ratio of time of play to advertising time than Premium service users.

- Based on the service plan selected, audio transmissions play advertising segments whose total length of time of play is determined according to a pre-defined ratio of advertising to airtime. For example, advertising may comprise 20%, 10%, or 0% of total airtime, according to the subscriber's service plan. In general, advertising segments are played in between tasks, so that the audio transmission is as smooth as possible. Lengthy tasks, however, may be interrupted by advertising. A task, clip, or procedure of no more than 2 minutes is immediately followed by advertising items with a total paying-time not to exceed the predefined ratio of play-to-advertising. (e.g., 2 minutes of audio transmission is followed by 12 seconds of advertising) Longer tasks are interrupted at the most convenient possible interval. All audio transmissions in basic service plan begin with a five-second advertising clip.
- The ratio of broadcast time to advertising time may be configured on a sliding scale such that the more active time a subscriber logs with the system, the fewer advertising items she hears. That is, the advertising algorithm is such that as the airtime increases, the ration of advertising to airtime decreases. In this way users are "rewarded" for increased airtime.
- Incentives may be offered to users in order to persuade users to listen to advertisements without interruption, or to persuade users to participate in surveys. Incentives include:
- (1) Discounts on system services
 - (2) Cash credit and vouchers for system "Get Me" services
 - (3) Discounts on commercial goods and services,
 - (4) Gifts to subscribers.
 - (5) Frequent-flier miles,

The system monitors subscribers' advertising listening habits, such that the appropriate rewards may be given, and the appropriate incentives offered to each user. Advertising revenues may be used to fund or subsidize program costs.

In the preferred embodiment, each segment of the audio transmission is coded to support an advertising item of a specified length. Articles are also coded for content, such that a sports article supports advertising items related only to topics of sports, health, and fitness.

In some embodiments, interactive advertising services are provided. Interactive advertising items may include surveys or providing coupons to users, and enable purchase during the advertisement.

Advertisements are segmented and woven into audio transmissions.

5.3 Description of Mobile Agent Applications:

5.3.1 *CELLCAST*

a Information editor is used to select stories and information items, and to edit and format these items such that they are represented in a form suitable for dissemination to users using the present invention. The selected and edited stories are then voice-recorded and stored in an information database on the Mobile Agent system.

In the illustrated preferred embodiment, the information editor categorizes each information item according to a predetermined set of criteria. The information editor maintains a list of currently defined categories and sub-categories. The personnel operating the Mobile Agent system may add and delete categories and sub-categories so as to accommodate major media events or special features. The list of category definitions is thus relatively constant but subject to change.

As new information items are received they are each assigned a weight relevance value of 1-100 against each category structure by the system editing crew. As the

- information items are accumulated, they are ranked based on the assigned relevance values. A cutoff threshold determined for each category structure is then applied to information items with respect to each category structure. If the relevance value for an information item exceeds the cutoff threshold for a given category structure (e.g., 60 points), a pointer identifying the information item is included in the category structure. The cutoff threshold is different for different categories and is generally empirically determined. As a result of the above operations, the system maintains a ranked list of information items received for each category structure.
- 10 A full sequence of assembling operations may be successively repeated (e.g., daily for a daily Cellcast report). Fig. 3, briefly referred to earlier, illustrates and example of the process for generating Cellcast summary reports. Generally, the individual assembling operations are performed for each profile. Since each profile may include a different scheme of prioritization, the relevance values are
- 15 separately tailored for each profile. These adjusted values are used to re-rank the information items for each category structure in the profile. The information items are then selected based on a priority scheme to create the final set.

The information database stores three recorded versions of each story. These are:

- *Headline*—A titular or nominal version of approximately 3-15 words, used primarily for organizing purposes;
 - *Short*—A executive summary version of approximately 15- 75 words; and
 - *Comprehensive*—A full story version of approximately 75 or more words, which is the most detailed version of the story that the present system
- 25 represents.

The information database includes statistics relating to the user's listening habits, storing information pertaining to the number of times any given item has been played, and the duration of each audio transmission. The database also includes

advertising play statistics indicating how many times each advertisement has been played to each user.

News, Features, Entertainment, Music

In the preferred embodiment, the information items selected for Cellcast play
5 consist of news and feature articles, entertainment features, music, and audio books.

In the preferred embodiment, Cellcast provides a wide range of musical selections to users, which are played on demand. Users may program personalized “radio shows” by choosing a play program for music selections. Cellcast also provides
10 preset play programs. Music is organized according to channels by genre: jazz, rock, classical, country, local or regional music, and so forth. Some items are available at a paid premium only.

In the preferred embodiment, Cellcast also provides audio books for users to listen to. Users may listen to audio books in segments of a predetermined length, or
15 may stop and start play as desired. Some items are available at a paid premium only.

Setting User Preferences for Cellcast

Personalization of services is via the Internet, or via cellular handset, depending on each user’s preference. The user defines the profile of audio content that may
20 be of interest to the user (e.g., specific stock quotes, or scores of specific sport teams.) This information is stored as part of the user-profile information database.

In the preferred embodiment, users select categories and sub-categories of information items. Each category may be set according to user preferences. Users determine:

- 25
- For which broad categories of information or “channels” users would like to receive and not receive data, within a predefined limit (e.g., user requests a sports channel as one of fifteen selected channels)

- Which sub-categories of data and information items are to be included and not included within each “channel,” also within predefined limits (i.e., user requests all stories of the football sub-category and no stories of basketball sub-category.)
- 5 • Which categories and sub-categories are to be included and not included as part of the user’s personalized default audio transmission. Users may also determine a maximum time allotment for their personal audio transmission.
- Which special features, programs, information items or data are to be included in the audio transmission. Users may direct the application to locate
- 10 special features, programs, information items or data using a search function.
- This process begins with each subscriber creating a profile by choosing relevant category structures (their “primary category structures”) for their own interests. Each subscriber determines the maximum number of information items and the maximum number of articles they want to receive each day. Next the process
- 15 continues with the system determining “secondary category structures” and “neighboring category structures” for use by the system on days when record volume is low (e.g., a slow news day) as received from information providers. Secondary category structures are user defined lower priority categories for a user’s profile, and neighboring categories are system-defined categories of related
- 20 subject matter. Both contain records that, while not of primary interest to the subscriber, are still relevant to the subscriber. Finally, the process distributes information items according to the limits set by the subscriber and the availability of information items in each of the primary, secondary, and neighboring category structures.
- 25 Each of the category managers includes a profiler procedure for defining the subscriber’s interest in receiving news items within each information category. For example, an “all” command can be used to select all sub-categories, and a “none” command can be used to indicate that the subscriber does not want to receive any news items for any given category. The category manager profile
- 30 procedure generates a category profile data structure that represents the

subcategories of interest to the subscriber as well as any associated filters that have been defined.

There are several different possible priority schemes that may be used to select the records to be included in the final set (i.e., in Cellcast reports). In the preferred embodiment Configuration for audio transmissions may occur at any time, including during the course of any given audio transmission. Any preferences that have been set automatically or that function as the mode of operations by default may be overridden with speech commands.

Users may pause the audio transmission at any point, and continue play when they so designate. Similarly, if the user hangs up in the middle of an audio transmission, she may return to the same point in the audio transmission when she calls back.

Personal Summary Report

In the preferred embodiment, the Personal Summary report represents the set of data and information items to be received by the user upon activation of Report mode, either by spoken command, or set as default. The personal summary report is read as a continuous audio transmission but is in fact comprised of a series short audio files of articles played in sequence. The stories are punctuated by a short pause after each story. Commands uttered during these pauses are to apply to the previous story.

In the preferred embodiment, the individual user determines the content, length, and style of the report.

In the summary report, each of the user's pre-selected channels is represented by an average of three stories, subject to user preference. Users may opt to exclude any selected channel from the standard summary reports, or to reduce or increase the number of stories to be included in any given summary report.

As stated previously, each short story corresponds to a “full story,” a longer and more detailed version of the short story, as well as a shorter “title” version, which is used mainly for sorting, browsing, and organizational purposes. In alternate embodiments, users may choose to include either of these versions in the personal summary report. The user may access these corresponding versions of stories by uttering the appropriate commands even when the summary report is being played. For example, during the summary report a short version of the audio report regarding a basketball match is transmitted to the user. During the transmission the user utters the command “full Story”. The system immediately starts transmitting the “Full Story” version of that same short story.

The matching software is responsible for matching the items according to relevancy and other record and user properties. The user profile and preferences are dynamically updated by the system. Tuning and redefining subscriber profiles is based on the subscriber’s usage feedback, which is developed by tracking the data requests issued by the subscriber, and usage statistics. In this manner, the usage feedback acts as an implicit, non-intrusive way for subscribers to let the system know which types of records they consider the most relevant. By ordering certain versions of an article, a subscriber is implicitly stating the relevance of that record to his or her interests. When several records of the same type (i.e. from the same category) are ordered, the statement of that category’s relevance to the subscriber becomes that more powerful. If the particular category in question has originally been placed by the subscriber low in the profile priority, the automatic profile tuning and redefinition process of this invention raises the category structure in priority to give it more prominence in the records or briefs delivered each day.

Reports may also include search results or specially requested programs, features, or information items.

Article Browser

In the preferred embodiment, the article browser is a program for listening to news items that the user specifically wants to hear. The browser can be launched at the user's explicit command. It can also be launched from the personal
5 summary report with the appropriate user commands such as, "full story," or "more details," indicating that the user wants to hear the full version of story. In addition, using the article browser, the user may hear headlines – the shortest recorded versions of each story—and may use appropriate commands to hear fuller versions, including both the executive version used in the summary report,
10 or the comprehensive version.

Users may either use voice commands to navigate between categories, sub-categories, and individual information items, or may direct the program to function in continuous mode within a given category or sub-category.

In the preferred embodiment, users may listen to any category or subcategory of
15 data or information items using the continuous mode. This is an option to hear news stories or data read in sequence, one item after the other, without requiring a new voice commands before each story is read. The continuous mode functions only within each given story category or subcategory and not between categories.

Thus users may command: "Go to CNN and read titles in continuous mode. Then
20 go to Washington Post and read executive summaries in continuous mode."

Premium Services

In the preferred embodiment, a Premium information retrieval service is available for business users and professionals. The Cellcast media team locates information on specialized topics of interest, customized to suit each user's field or specialty.
25 Users may request news, information, press releases, research, reports, and reviews on a vast array of topics. For example, a user may request information about mergers and acquisitions in the telecommunications industry in Europe, data relating to recent fluctuations in crude oil prices in the Middle East, or reviews of recently released books on Internet-related law.

Premium service uses the broadest possible range of Internet-based sources in order to obtain the most pertinent information, including news feeds from a number of information transmission services, hundreds of information databases, and full access to the World Wide Web.

- 5 A computerized tagging, mapping, and filtering system combines forces with a human editing team to determine which sources are the most accurate and relevant to the users' chosen topic(s). Articles are prioritized and categorized according to relevance and user preference.

- 10 Premium service reports are fully integrated with all other Cellcast functions and properties.

- Premium service users may also track a specified set of named entities, such as companies. Embodiments of the retrieval system of the invention can include a process whereby the subscriber selects a collection of records containing relevant information about any of a specified set of named entities from a larger set of
- 15 records whose content may be either relevant or non-relevant to the set. The relevant information can include the full set of information items relevant to the companies or named entities in the set, or a subset of those records determined by additional subject matter criteria.

- The tracking process includes a multi-stage, rule-based system that attaches one or
- 20 more tags to an information item corresponding to each company or named entity specified as a member of the set. The information items are collected and rule-base tags are attached to them corresponding to each company or named entity that is part of the specified set. Information items are then sorted and evaluated accordingly. The resultant set of information items may be played as part of the
- 25 subscriber's summary report, or stored on the subscriber's personal hard disk for future reference.

Personal Research Assistant

In the illustrated preferred embodiment, users may access a Personal Research Assistant function via the application server 6. This function allows the user to obtain answers to specific queries in addition to a compilation of articles on relevant topics. The Personal Research function combines the specialization and broad-based information access of the Premium service with full human capabilities.

Personal Research Assistant functions are tailored to high-level, complex queries. An example of such a query might be, "How does the real estate market in California change directly after earthquakes," or "How has the quality of women's lives in Afghanistan changed over the last ten years?"

Answers to questions can be represented by special reports, recorded specifically in response to the user's question; a series of articles edited and filtered by a human research assistant; or a combination thereof.

The range of topics available to users is unlimited. Human editors and researchers work in conjunction with Cellcast's broad information retrieval network, checking the results of all searches and personally supervising the research, editing, and recording process.

Besides searching all available databases, the Personal Research Assistant will locate the requested information when data is otherwise unavailable.

5.3.2 General Information Services Function

In the preferred embodiment, user-initiated queries are answered via the cellular network. Users access a selection of short-answer tools through the Mobile Agent, which are together classified as General Information Services of the application server 6. These services include, but are not limited to, reference information, tools for calculation, and short entertainment features. The General Information Services are designed to obtain information, and to perform specific functions, as well.

Natural speech queries are converted to the required data forms using advanced query applications and a natural language engine. Based on the data form content, the Mobile Agent system database is searched for the required action, process, or information item. The user's request is duly processed, and the results of the query are relayed to the user in an audio format.

Users initiate queries by asking questions or by requesting the appropriate tool. Requesting a specific tool serves as a cue to the system as to which operations are to be executed using the user-input data. For example, if a user wants to convert five dollars to yen, the user may ask for the exchange rate using natural language, or alternatively, the user may access the currency exchange tool and then directly convert the sum.

In the preferred embodiment, the General Information Services include but are not limited to the following services:

- A dictionary, a thesaurus, translation tools and grammar tools;
- A calculator, calendar information, clock information, weather information, currency exchange, and conversion tables;
- People search, directory assistance, yellow pages, and white pages;
- Restaurant, television and movie guides;
- Games;
- Horoscopes;
- A diet manager application;
- A medical information guide;
- Guides for auto purchase, job hunting, and house hunting; and
- Classified advertisements.

A portion of the General Information service functions are location sensitive, providing information and services particular to the user's location. Yellow and white page directories, TV and movie listings, and classified advertisements are

among the information categories that are organized according to the user's location.

5.3.3 Broadcast Media Hyperlinks

The system enables phone users to interact with traditionally two-dimensional and one-way media sources, such as radio, television and print media. Just as the Internet allows browsers to retrieve additional and associated information through the hyperlink, cellular media hyperlinks allow phone users to hyperlink to radio and TV broadcasts and billboard and magazine adverts. Thus, radio listeners will be able to buy any song that was recently played, or find out more information about any product, which they have just heard advertised, and interactive radio commercials are made possible. The system includes hyperlinks relevant to each of the broadcast tracks.

Upon hearing a song or advert which s/he liked, the listener will access a CellCAST's WAP or Web site, and select recent item type the user is interested in Music, Commercials , or News.

the user is then presented with the recent music tracks or commercials played on the radio, and obtain additional information on the item or buy a related product (e.g. music CD). In the case of a commercial, Cellcast will be able to connect the listener directly with the advertiser.

When commercials are played on the radio, advertisers who take advantage of Cellcast will be able to "stretch" their 30-second radio spot, and provide extra information for interested listeners. This direct connectivity will also allow advertisers to measure the effectiveness and performance of their commercials.

5.3.4 "Get Me" Active Transaction Services

a Get Me application and database for ordering goods and services including information items, media programs, and other commercial products according to the respective command; The system functions as an "agent" empowered to buy products according to the user's commands. These operations may be executed according to a profile of pre-defined user-preferences or system options, and in online, real-time transactions from radio stations/TV, etc. Via the cellular network users may access travel agent services, order a pizza using only a cellular phone,

buy a product that is being advertised on the radio or TV, or buy in-depth an Interview that was broadcasted on the radio or TV.

In the preferred embodiment users may make purchases and conduct financial transactions via the “Get Me” services of the application server 6. The Companion
5 functions as an agent empowered to buy and sell products according to the user’s commands. “Get me” services are usually but not necessarily commercially oriented; some “get me” functions do not involve financial transactions. For example, a subscriber may request, “Get me a copy of tonight’s news audio transmission.” The system might be able to provide a video-file of the news free
10 of charge. A subscriber may request a copy of today’s Howard Stern radio show, or a copy of the song played on the show, or one of the items mentioned in the show or the advertisement played on the show.

A natural speech engine in the appropriate speech recognition server 3 processes users’ requests. Human voice commands are converted to data forms recognized
15 by the system, and the users’ command is then extracted and processed accordingly. The task is executed after user confirmation without further human intervention.

Linked to PCA calling functions, “Get me” services may connect users to sales representatives instead making purchases. Users may choose to grant agency to
20 the system, or may choose to be conduct transactions personally after having been connected to the appropriate number.

In the preferred embodiment, users are billed for purchases and transactions via the cellular operator. In an alternate embodiment, users are billed via the Credit Company of the user’s choice.

25 User need not be on the line in order for transactions to take effect.

“Get me” operations may be executed according to a profile of pre-defined user-preferences or system options. For example, if a user request, “get me a pizza,”

the system may “know” to order from a specific establishment according to the user’s preset preferences. Alternatively, the system may “remember” that the user always orders pizza from a specific establishment, and will proceed with the operation accordingly.

- 5 Advertising items linked to this function may offer services that put the means of carrying out the transaction at the user’s immediate disposal. For example, if the user requests, “get me a pizza,” the user hears an advertisement suggesting that the order be from a particular establishment. If the appropriate voice commands are used, the operation will be executed at once.
- 10 Possible transactions include, but are not limited to the following:
- Online travel agent services
 - Recording television programs upon user request
 - Online trading
 - Taxi service
 - 15 • Movie tickets
 - Banking/payment services
 - Any other available online shopping or interactive service via Internet

5.3.5 Remote access to PC via cellular phone

- 20 As shown in Fig. 4, access to the information stored on the Virtual PC is available either by connecting to the PC directly using the mobile phone, or by accessing synchronized copies of information, data, or files residing on the system servers.

Files are synchronized with those on the user’s own personal computer 412, and server-side application settings may be synchronized with those on the user’s PC

25 413, as well.

The following synchronization options are available:

- (1) files are copied or upon request modified via dial-up Internet connection;
- (2) synchronization occurs according to a pre-defined schedule (e.g., every two hours) using dial up internet connection, leased line, frame relay,
5 or any other form of internet connectivity;
- (3) synchronization occurs immediately via networked Internet connection.

Synchronization is available in both directions:

- (1) subscribers synchronize files and directories on their own personal
10 computers to correspond to Virtual PC applications.
 - (2) subscribers synchronize Virtual PC or server-side files to correspond to files on their personal computers. files in the specifies directories on the user's PC are copied to the system server and saved; or copies of all user-designated files are saved to the user's personal computer.
- 15 If the user's personal computer is connected to the Internet directly or directly to the cellular switch, files on that computer may be accessed directly by the user, and synchronization is not required.

The Virtual PC allows mobile users to access files and run applications on the server-side using files from the user's personal computer, once synchronization
20 has occurred. Access to the synchronized file servers is made via the cellular phone and the mobile Agent system or via the web directly to the Mobile Agent file servers.

The application server platform enables the addition of third-party applications, so that they may be run or accessed using the cellular handset via the Mobile Agent.

- 25 Users may access any information item by category, specific properties, name or location, and by topic. Users create their own directories and sub-directories for storing and organizing information (see FIG 7). Users can receive specifically requested articles or other information items directly into specified folders.

Information is stored in the user database as pointers to stories on the in the relevant database. Information includes personal data (contacts, appointments) saved pointers to Cellcast articles, saved pointers to Road directions, saved answers to user-initiated queries, and the results of user-initiated searches. Each
5 separate information item is tagged, coded, sorted, and stored on the system database.

5.3.6 *Internet Access and Web Searches*

Upon accessing the world wide web via the system the user reach a portal of VXML compliant web sites. Sites using VXML protocol can be played via the
10 seamlessly operating natural language engine 4. Access to other sites is enabled using a text to speech engine of speech recognition server 4 and converting the text to synthesized digital audio signals, which are sent to a mobile station via the cellular network.

Web interface includes the following options:

- 15 1. "Browse" the Mobile Agent indices of web sites by topic heading. Indices of available topics are read as lists upon user request. Users may browse categories and subcategories of web sites rather than conduct a search.

Set a search as a default mode of operations. Thus, a user seeking stories, news,
20 or information items about the Washington Redskins may conduct a search any time she is on the line; or, she may set a search as a default mode such that search results are a standard part of her personal audio transmissions.

In the preferred embodiment, users may access a search engine using voice commands. The search interface comprises of the following options:

- 25 (1) search the system database; (2) Search the World Wide Web; (3) Search only given category or sub-category of information on the system database, or World Wide Web, or from available media streams.

Searches may be conducted by keyword or by topic. If by keyword, files are searched for all occurrences of the specified word. If by topic, then files are

searched according to tags. Human editors assign the tags as part of the categorization and filtration process; all articles are tagged according to topic. Each article generally has a number of tags assigned to it. The Mobile Agent features book-marking capacities, tracks changes users make, records history, and
5 saves user favorites.

In an alternate embodiment, the same services are available using packet switching such as in GPRS systems whereby the user downloads data packets that are assembled at the user-side, thus comprising full data files. In such a system, saving information both on the user handset and at the switch would be possible,
10 and more efficient downloading of certain information would be possible sending information packets to the mobile handset while the user is in network cells with a cell capacity load not exceeding a predetermined threshold, thus not creating a burden on frequency reuse.

5.3.7 Road Direction Services

15 In the preferred embodiment users may obtain location-sensitive road directions to and from any point (within designated areas) by connecting to the road direction application of the application server 6 and the relevant database. Directions are comprised of a series of incremental pre-recorded audio segments corresponding to segments of the selected route, linked together sequentially.

20 Road directions are based on the user's present location and the user's desired destination. Using this data, a search is automatically conducted for directions.

In the preferred embodiment, users input their present location using simple speech commands, following voice-prompts. The user's location information is located on the server's map and cross-referenced with a database on the Mobile
25 Agent system. If any information is incomplete or inconsistent, the user will be prompted to provide additional details relating to the present location. The user's location is determined using a combination of voice-input by the user and network information. For example, the user may indicate that he is presently located on

Main Street, and the network information services will indicate in which city this particular Main Street is located.

In an alternate embodiment, the cellular network may determine the user's present location using GPS reports or any other tracking systems or techniques such as
5 triangulation.

Once the user's location is verified, the user enters the desired destination information, still using voice prompts. Users enter specific destination information, such as a specific street address, or general destination information, such as "the nearest hospital." Users may enter multiple destinations, and may
10 specify an order in which to reach those destinations.

Directions are suitable for pedestrians as well as drivers. Information may be customized to suit drivers of vehicles with special requirements, such as trucks, vehicles carrying hazardous materials, or lightweight vehicles such as bicycles and mopeds. Directions for using available public transportation systems are
15 available, as well, including relevant timetables. Road directions may accommodate combinations of various forms of transportation, such as walking, biking, and subway riding, for example.

In the preferred embodiment, the road direction function is linked to a constantly updating traffic information database. Users' requests for directions are
20 accompanied by relevant information relating to traffic congestion, road construction, and driving conditions.

Once directions are found and compiled, users may listen to directions all at one time or in segments en route to the destination. The user may save road directions on his personal hard disk for future reference.

25 While listening to road directions en route to the destination, the user will be prompted to indicate when he has reached the end of a given segment. For example, the user is prompted, "Turn right on Main St. and drive 2.2 kilometers to

Pine St. There will be a hospital on your right. Tell me when you get there.”

When the user reaches the hospital he indicates as much. This functions as a command to proceed with the next segment of directions.

During pauses between segments, advertisements targeted to the user and relevant to the user's location are played. The advertisements are targeted to users based on their individual profiles, and are also targeted to users based on their location, destination, route, and mode of transportation. For example, a driver may hear an advertisement referring them to a car wash on his route, while a pedestrian may hear an advertisement for a restaurant on his way.

- 10 During pauses, news and information items requested by the user are also played. The user may also listen to road directions in conjunction with personal audio transmissions or while retrieving any other information.

- 15 In the preferred embodiment, the road direction function works in conjunction with a location information function. Linked to the road direction database, the location information feature allows users to obtain specific information related to either their present location or intended destination. Location information is linked to local directory assistance functions, as well. Location information includes but is not limited to the locations of: hospitals, pharmacies, restaurants, movie theatres, airports, train and bus stations, museums, and shopping facilities.

- 20 Road directions may be used to plan trips, to find alternate routes to clogged or congested ones, and in emergencies.

The flowchart of Figs. 8a, 8b illustrate an example of an overall operation involving accessing the system to obtain both personal cellular assistant (PCA) services and road direction services.

25 *5.3.8 Additional Features and Services*

Help Functions: In the preferred embodiment help services are available from any of the functions and services accessed via the Mobile Agent. Upon request for

help, users are told which commands are relevant or appropriate at present; users may request help for a specific topic; or users may receive explanations of specific features and functions. The help received depends on both the specific commands given by the user and the point in the audio transmission from which help is entered.

Help is automatically given to the user when accessing any service or function for the first time, unless the user chooses otherwise.

Human help is also available. Users may request the system help desk in order to access human help.

10 *Reports:* In the preferred embodiment, users may obtain reports of all activity conducted via the Mobile Agent. Reports contain all relevant information pertaining to Cellcast reports, general inquiries, road direction inquiries, Personal Cellular Assistant functions, and “get me” functions. Reports provide an accurate record of activities including but not limited to calls, transactions, and information
15 sent and received via the Mobile Agent.

Users may select categories and sub-categories of information to be included and not included in reports. Configuration of reports is available via the Internet using a PC or via the Mobile Agent.

Employers or corporate users may use reports to monitor employees’ use of the
20 system. All activities are stamped with time and date information. Reports may be delivered as text or as audio files.

As an example of a report, a user might request a textual report of all messages sent via the Mobile Agent over a period of 14 days, to be sent as an e-mail to the user.

25 **5.4 Alternate Embodiments and Extensions**

The above description is intended to illustrate the capabilities and operating procedures of the present invention and is not intended to delimit or restrict the

full range of capabilities made possible by the invention. A number of alternate embodiments and extensions are possible. For example, a personalized audio transmission identical in format and content to that which is available via the cellular network will be made available via the Internet, providing automatic
5 download of MP3 audio files to user's PC. Modifications and changes to the above-described invention are possible and permissible without departing from the essential spirit of the invention.

WHAT IS CLAIMED IS:

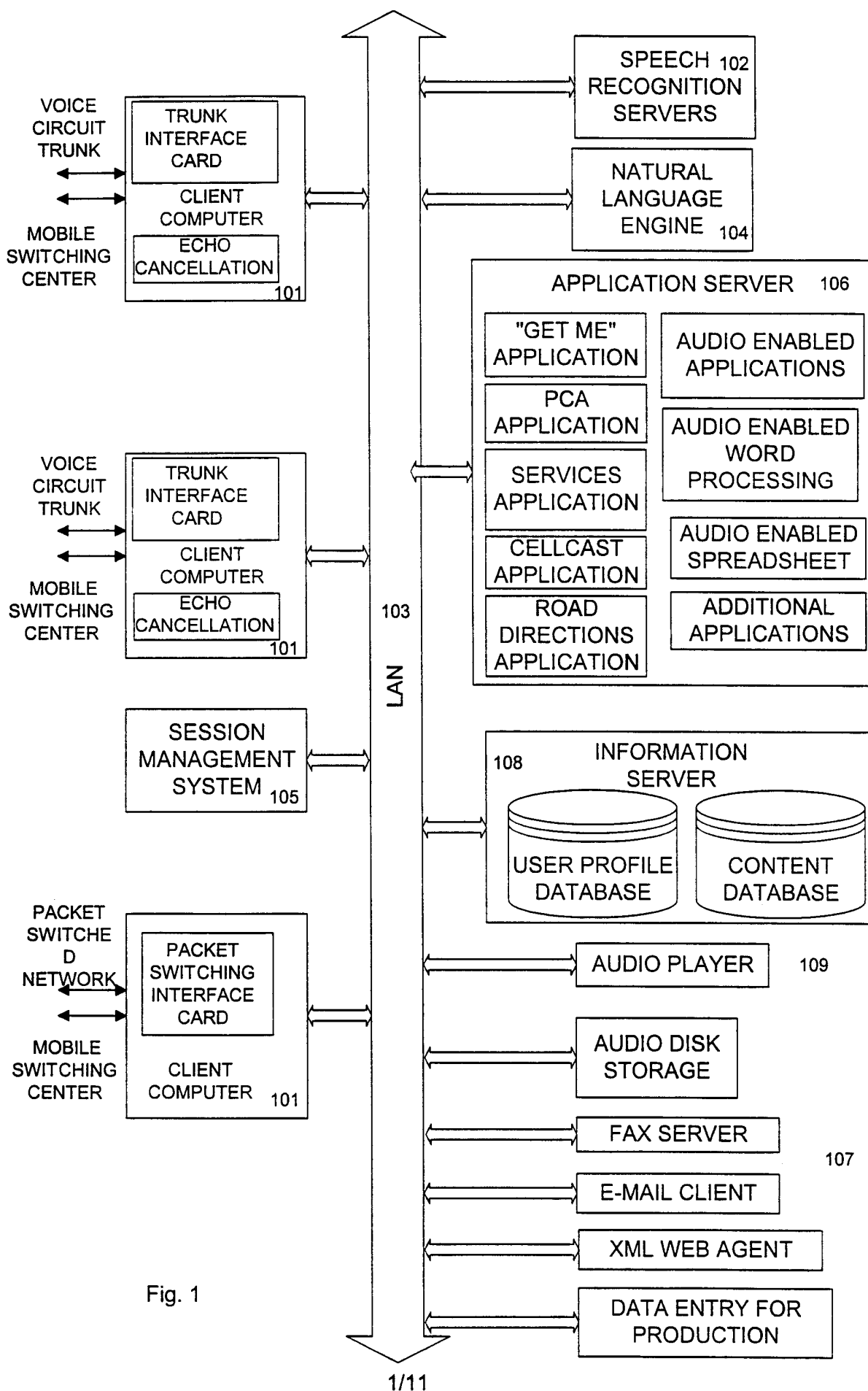
1. A platform enabling a plurality of cellular telephones access and interact with information using a speech interface over a telephony network comprising:
5 a recognition engine for converting speech received from the subscribers telephone handset to commands;
a speech interface including a speech commands based on command options information retrieved from a database and additional speech commands added to improve speech recognition and understanding;
10 and application software operated by the speech commands originating from the cellular telephone.
2. The system according to claim 1 wherein said speech interface identifies patterns of regular use according to predefined formulae, and
15 measures the occurrence of identified patterns; upon measuring recurrence of the pattern above a predefined level, the system adapts by setting new default modes of operations.
3. The system according to claim 1 wherein said speech interface can be presented to the user with different characteristics including prompt
20 voice attributes. Users can select a persona from a list of predefined settings resulting in different speech interface characteristics .
4. The system according to claim 1 wherein said application software enables access to the user's PC.
5. The system according to claim 1 wherein said application software
25 enables access to a virtual PC maintained for the user on the system database.

6. The system according to claim 1 wherein said application software enables ordering goods according to the respective command and the system functions is empowered to buy products according to the user's commands
- 5 7. The system according to claim 1 wherein said application software enables management and access to media hyperlinks
8. The system according to claim 1 wherein said application software enables ordering items referred to on the electronic media according to the respective command
- 10 9. The system according to claim 1 wherein said application software enables users to create media bookmarks according to the respective command
10. The system according to claim 1 wherein said application software enables ordering songs broadcasted on the electronic media according to the respective command
- 15 11. The system according to claim 1 wherein said application software enables ordering news stories broadcasted on the electronic media according to the respective command
12. The system according to claim 1 wherein said application software enables accessing information related to advertisements broadcasted on electronic media according to the respective command
- 20 13. The system according to Claim 1, wherein said application software enables road information dissemination according to the respective command.

14. The system according to Claim 1, wherein said application software enables a dissemination of news information to be supplied to the subscribers according to the respective command.
- 5 15. The system according to Claim 1, wherein said application software includes a General Information database containing reference information, tools for calculation, games, and other general information, to be supplied to subscribers according to the respective command.
- 10 16. The system according to Claim 1, wherein said application software enables each subscriber includes a Profile database for storing, updating, and retrieving personal data regarding preferences of the respective subscriber, and a Content database for storing data as required by the subscriber enabling such data to be retrieved as desired by the subscriber.
- 15 17. The system according to Claim 1, wherein the system further comprises interfaces for external devices, including an audio player, an audio disk storage, a fax server, and an e-mail client, all controlled by said session management system.
- 20 18. A Natural Language Engine which interacts with a recognition engine. The Natural Language Engine retrieves the data that is to be accessed with a speech interface, then implements several processing modules and prepares an improved speech interface.
19. The system according to claim 18, wherein said natural language includes recognition grammar caching.
- 25 20. The system according to claim 18, wherein said natural language enables adding suffixes.

21. The system according to claim 18, wherein said natural language enables adding alternative description of items on command option list.
22. The system according to claim 18, wherein said natural language enables multiple grammar fields in one sentence input
23. The system according to claim 18, wherein said natural language enables ambiguity detection.
24. The system according to claim 18, wherein said natural language enables enhanced navigation commands
25. The system according to claim 18, wherein said natural language enables foreign accent adaptation.
26. The system according to claim 18, wherein said natural language enables prompt enhancement.
27. The system according to Claim 18, wherein said server, in addition to enhancing speech recognition and grammar. also converts on-the-fly one standard markup language to another.
28. The system according to Claim 18, wherein said database also stores matching TTS prompts and natural prompts, and converts, on the fly, TTS prompts to its natural prompt. a text-to-speech compiler for converting text to speech;
29. The system according to Claim 18, wherein said server precompiles TTS prompts, and stores them in its database as recorded prompts
30. The system according to Claim 18, wherein said server automatically adds pauses and intonation tags according to said predetermined rules.

31. The system according to Claim 18, wherein said server automatically converts between different types of audio formats, grammar formats, and other extended formats according to the file type being processed.



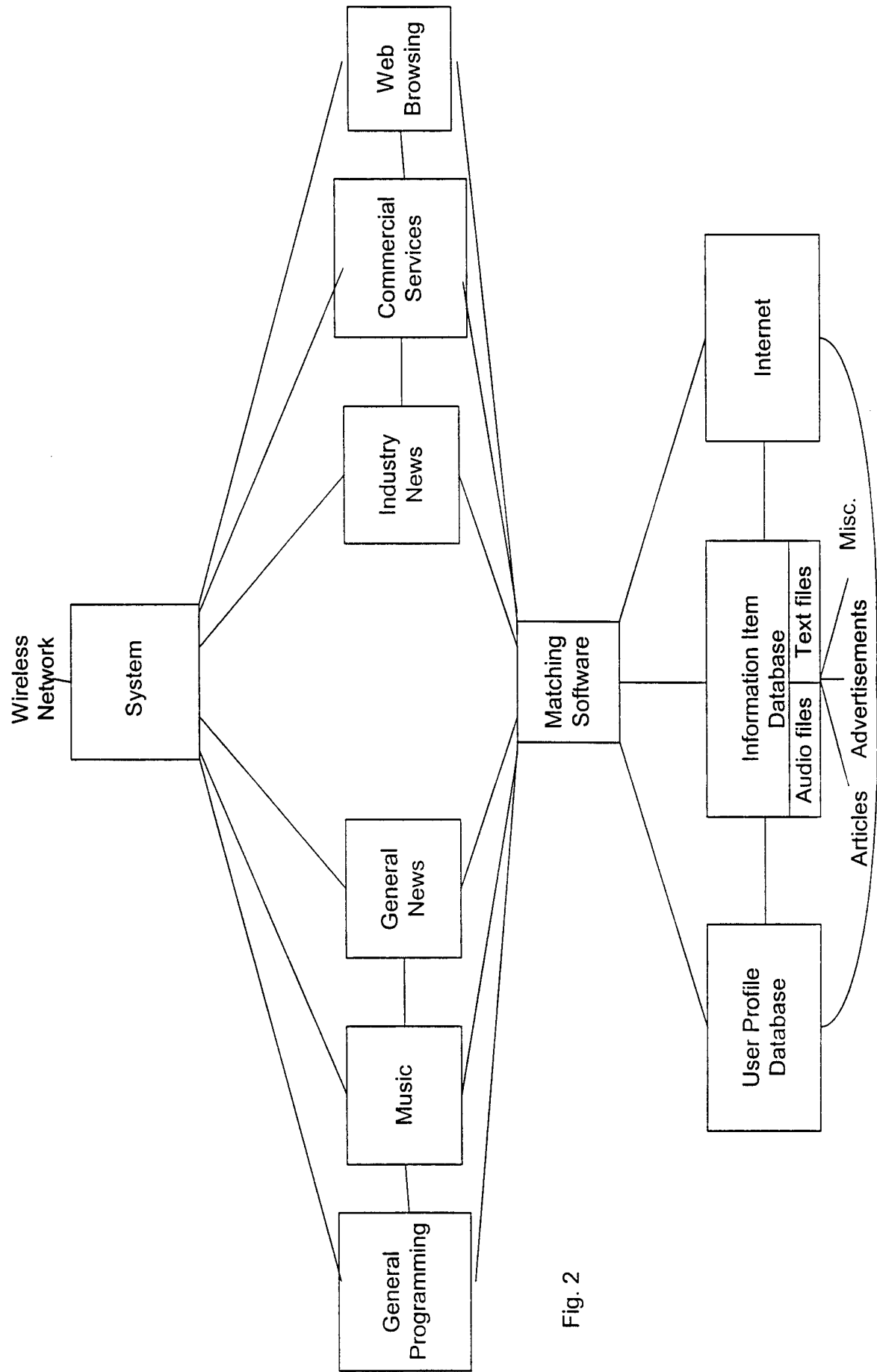


Fig. 2

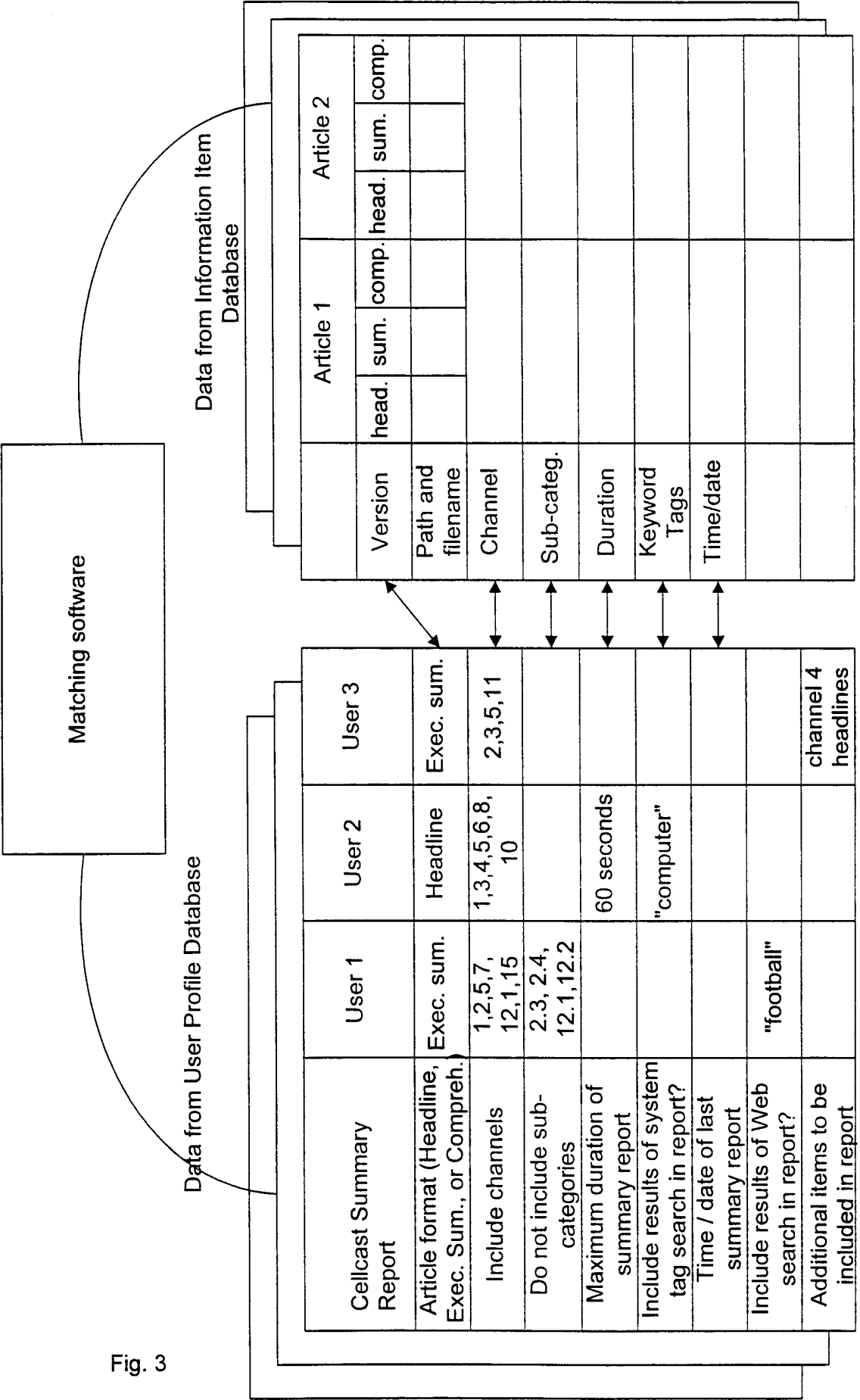


Fig. 3

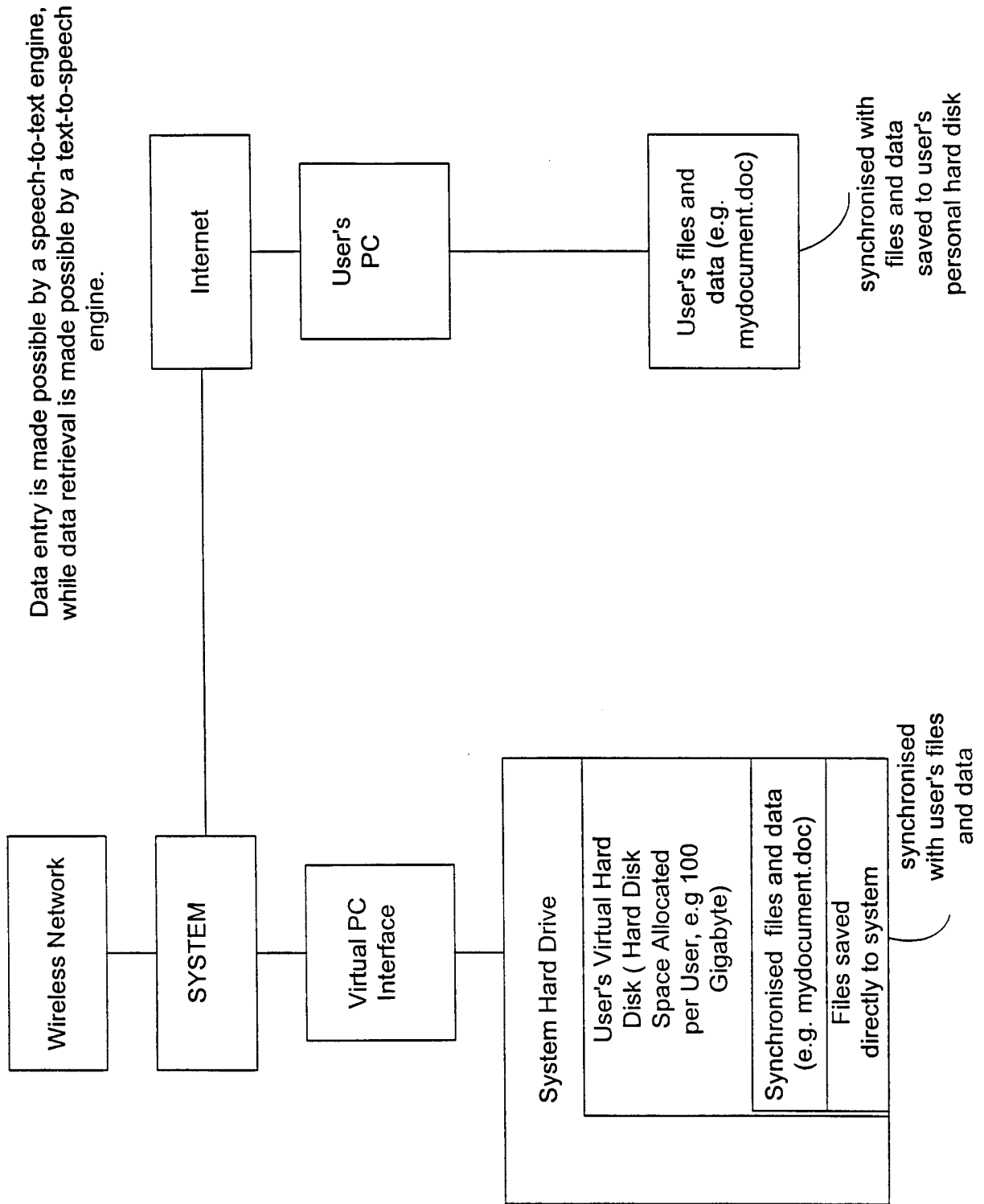


Fig. 4

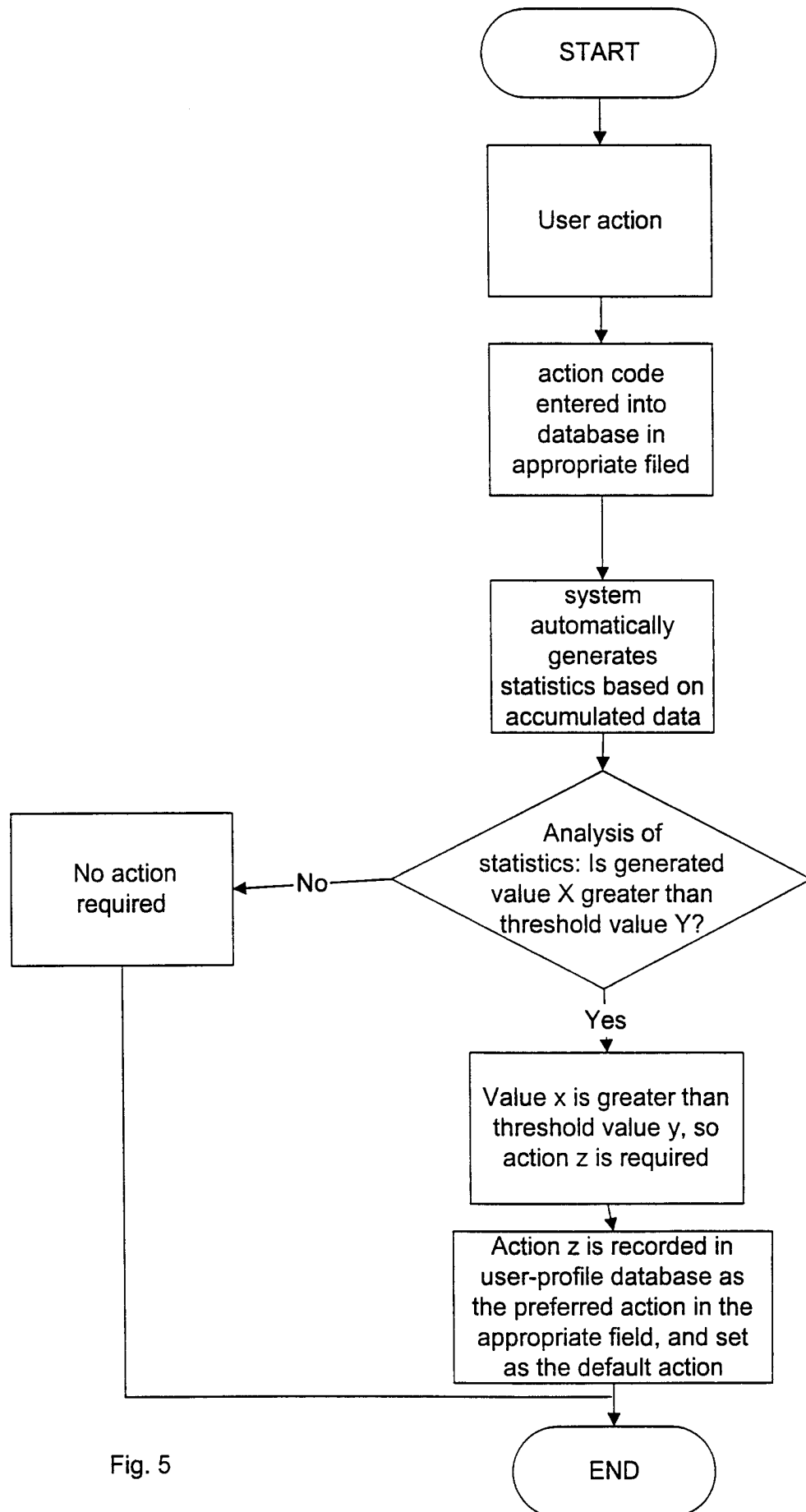


Fig. 5

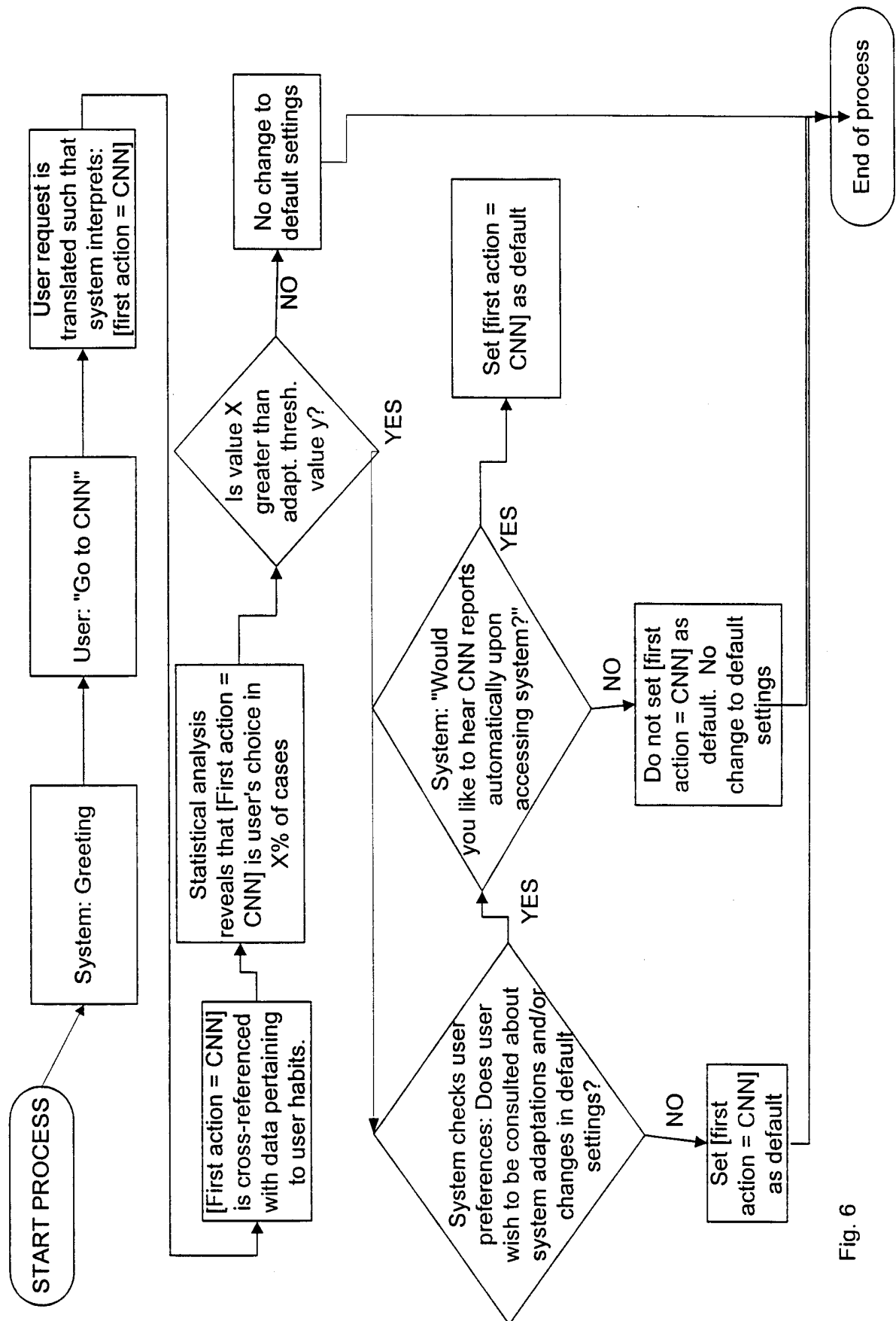


Fig. 6

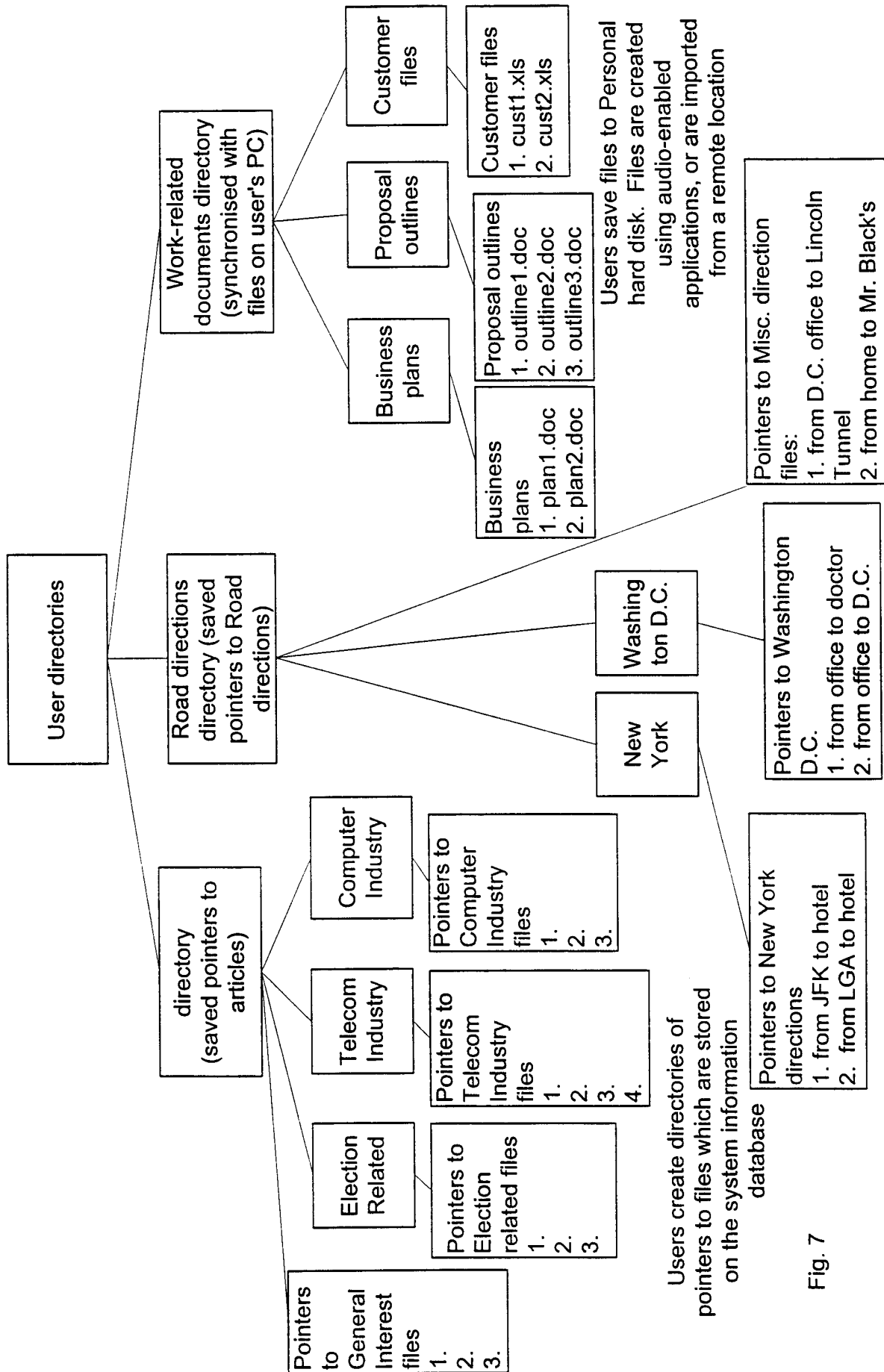


Fig. 7

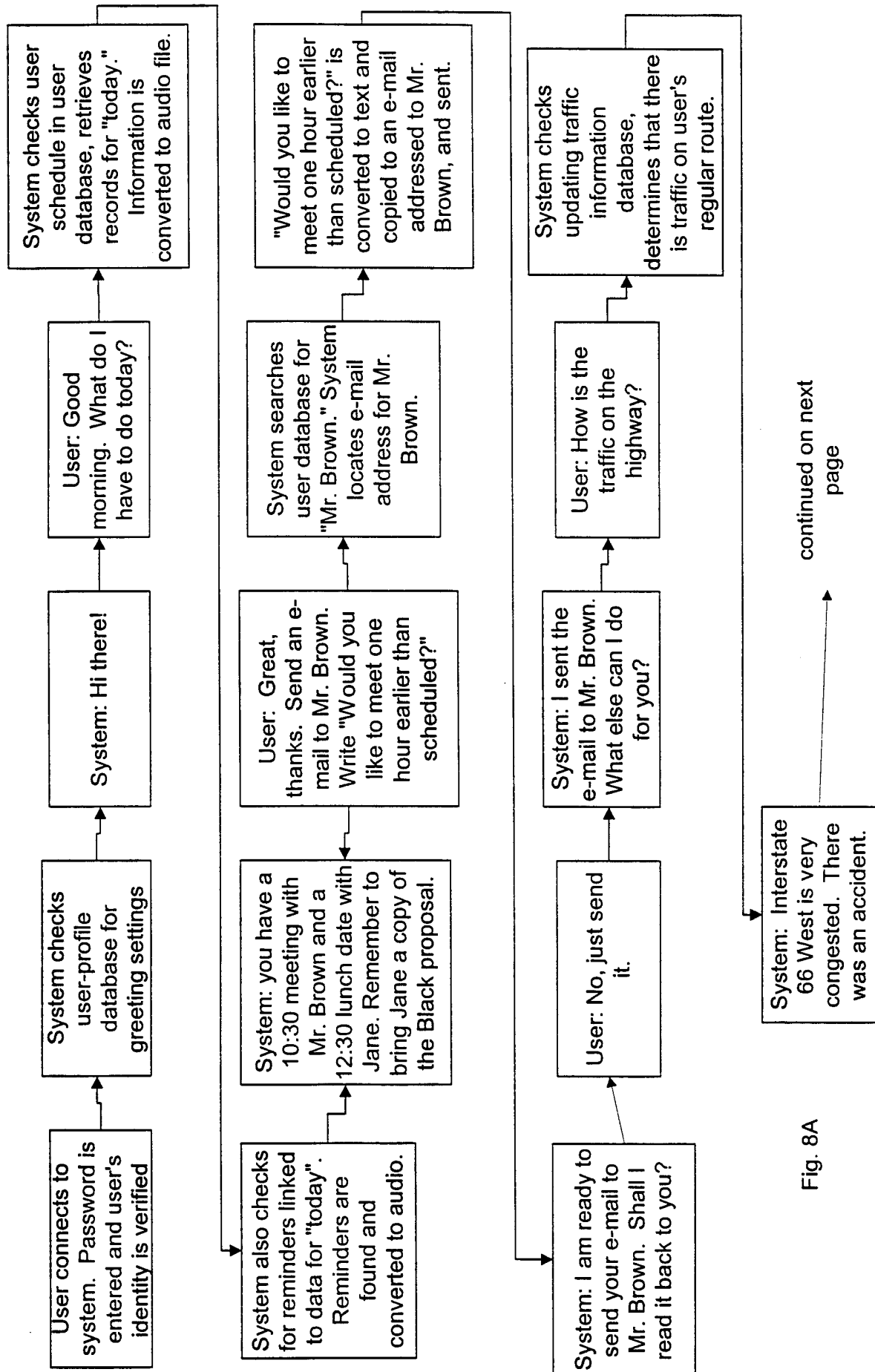


Fig. 8A

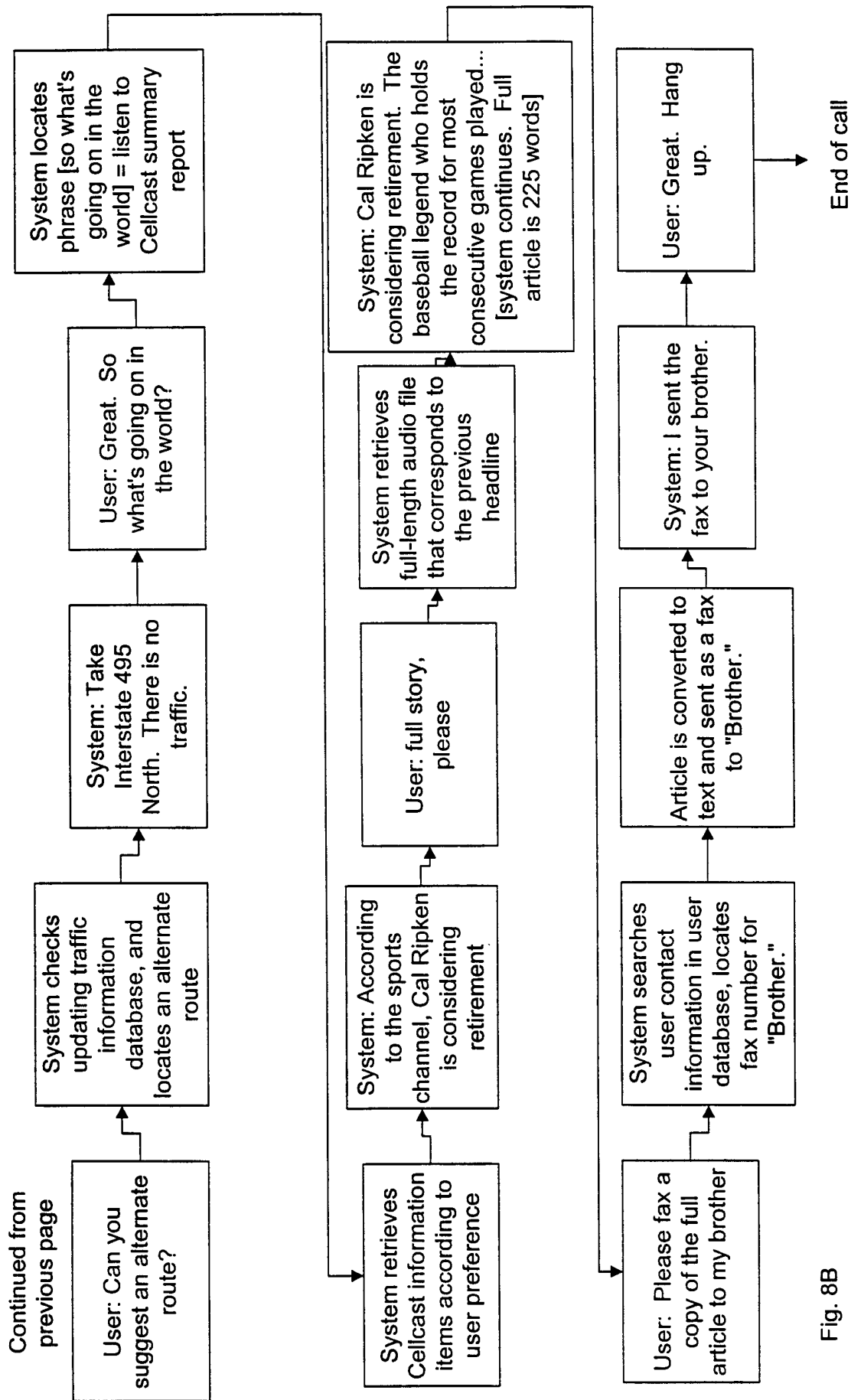


Fig. 8B

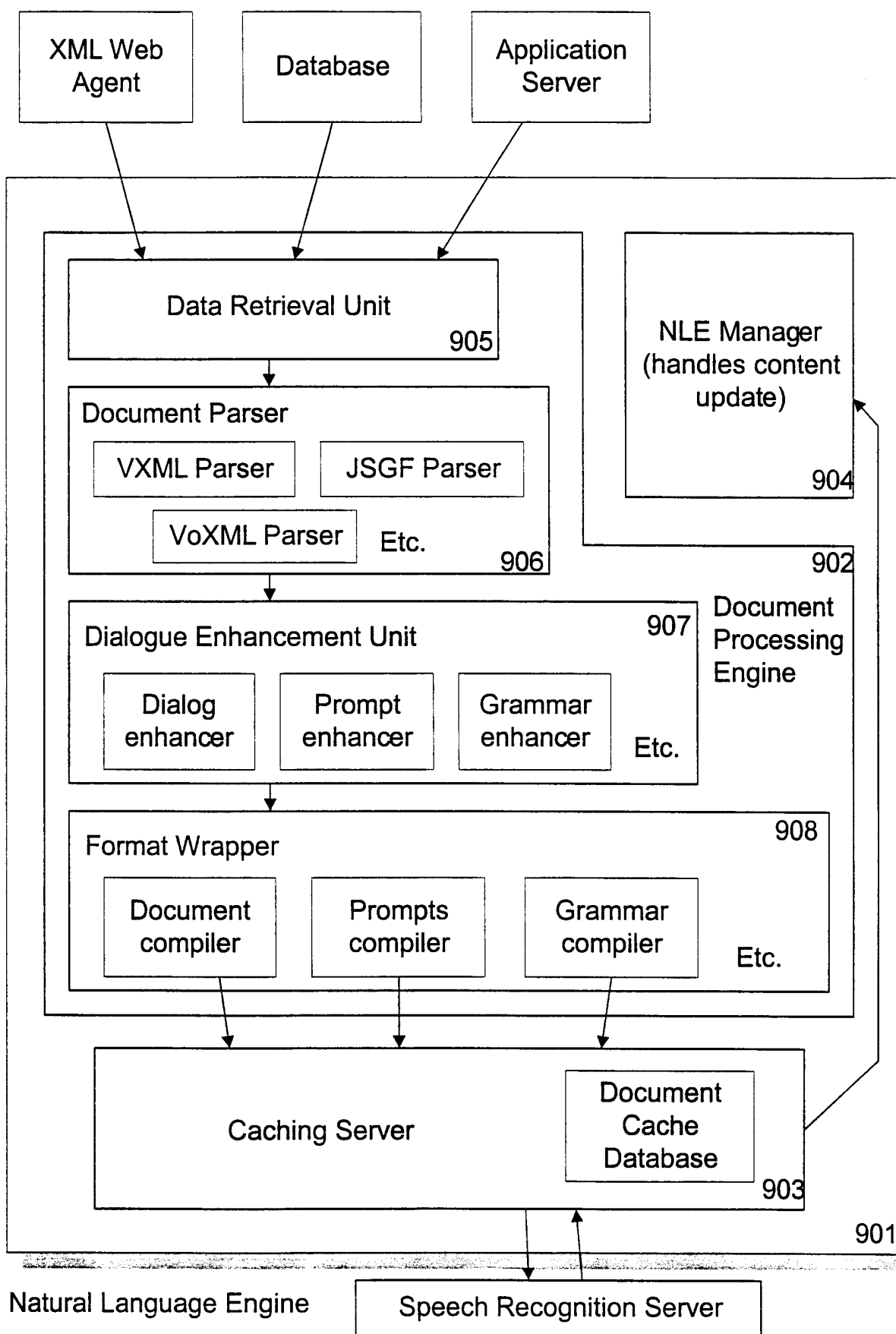


Fig. 9

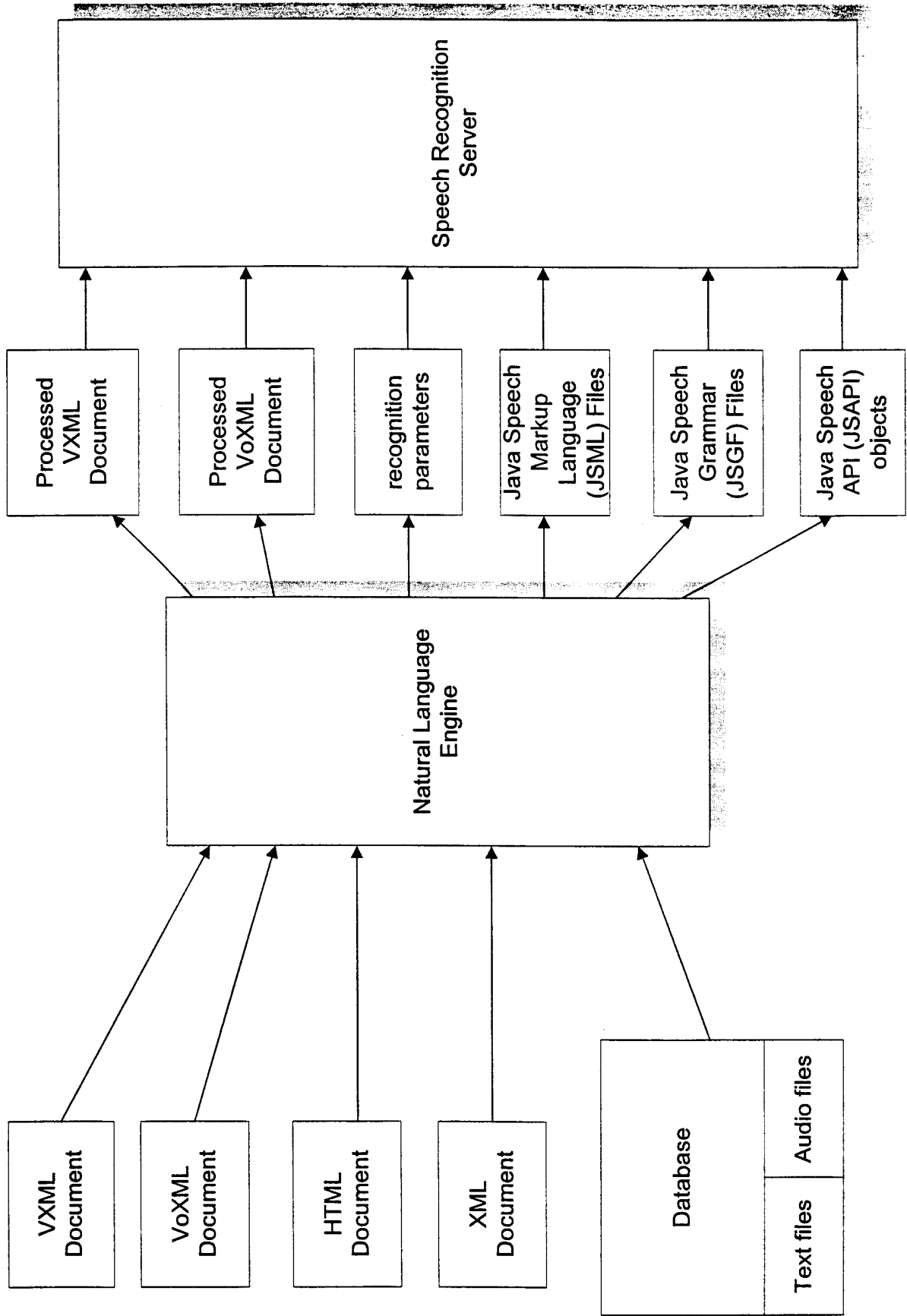


Fig. 10