US 20140282006A1

(19) United States

(12) Patent Application Publication

(10) Pub. No.: US 2014/0282006 A1

(43) Pub. Date: Sep. 18, 2014

Bolchini

(54) **AURAL NAVIGATION OF INFORMATION RICH VISUAL INTERFACES**

(71) Applicant: **Davide Bolchini**, Indianapolis, IN (US)

(72) Inventor: **Davide Bolchini**, Indianapolis, IN (US)

(21) Appl. No.: **14/024,612**

(22) Filed: **Sep. 11, 2013**

Related U.S. Application Data

(60) Provisional application No. 61/799,748, filed on Mar. 15, 2013.

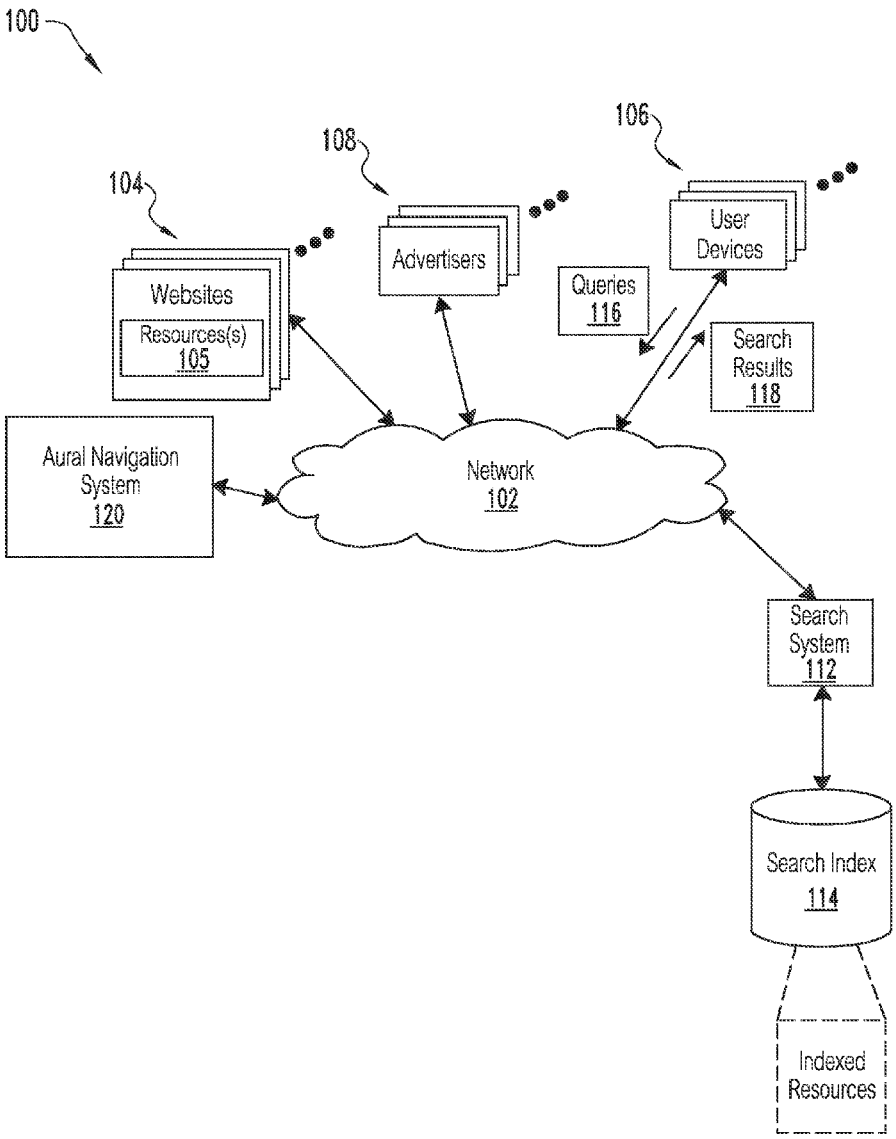
Publication Classification

(51) **Int. Cl.**
G06F 3/16 (2006.01)

(52) **U.S. Cl.**
CPC *G06F 3/167* (2013.01)
USPC **715/728**

(57) **ABSTRACT**

A method comprising generating, by a computer, a model of a website using user interaction primitives to represent hierarchical and hypertextual structures of the website; generating, by the computer, a linear aural flow of content of the website based upon the model and a set of user constraints; audibly presenting, by the computer, the linear aural flow of the content such that the linear aural flow of content is controlled through the use of user supplied primitives, wherein, the linear aural flow can be turned into a dynamic aural flow based upon the user supplied primitives.



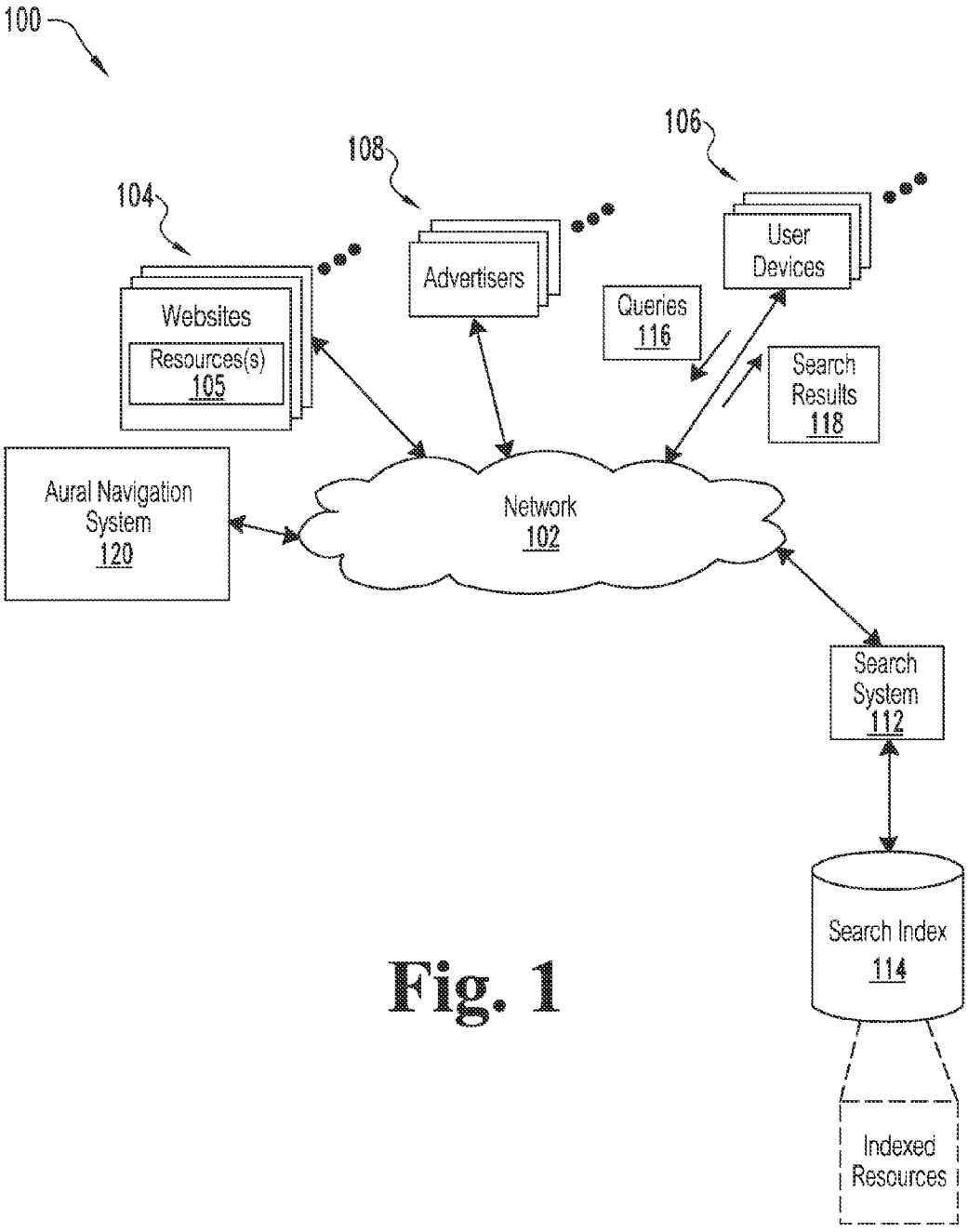

100
104
108
106
Websites
Resources(s)
105
Advertisers
User
Devices
Queries
116
Search
Results
118
Aural Navigation
System
120
Network
102
Search
System
112
Search Index
114
Indexed
Resources

Fig. 1

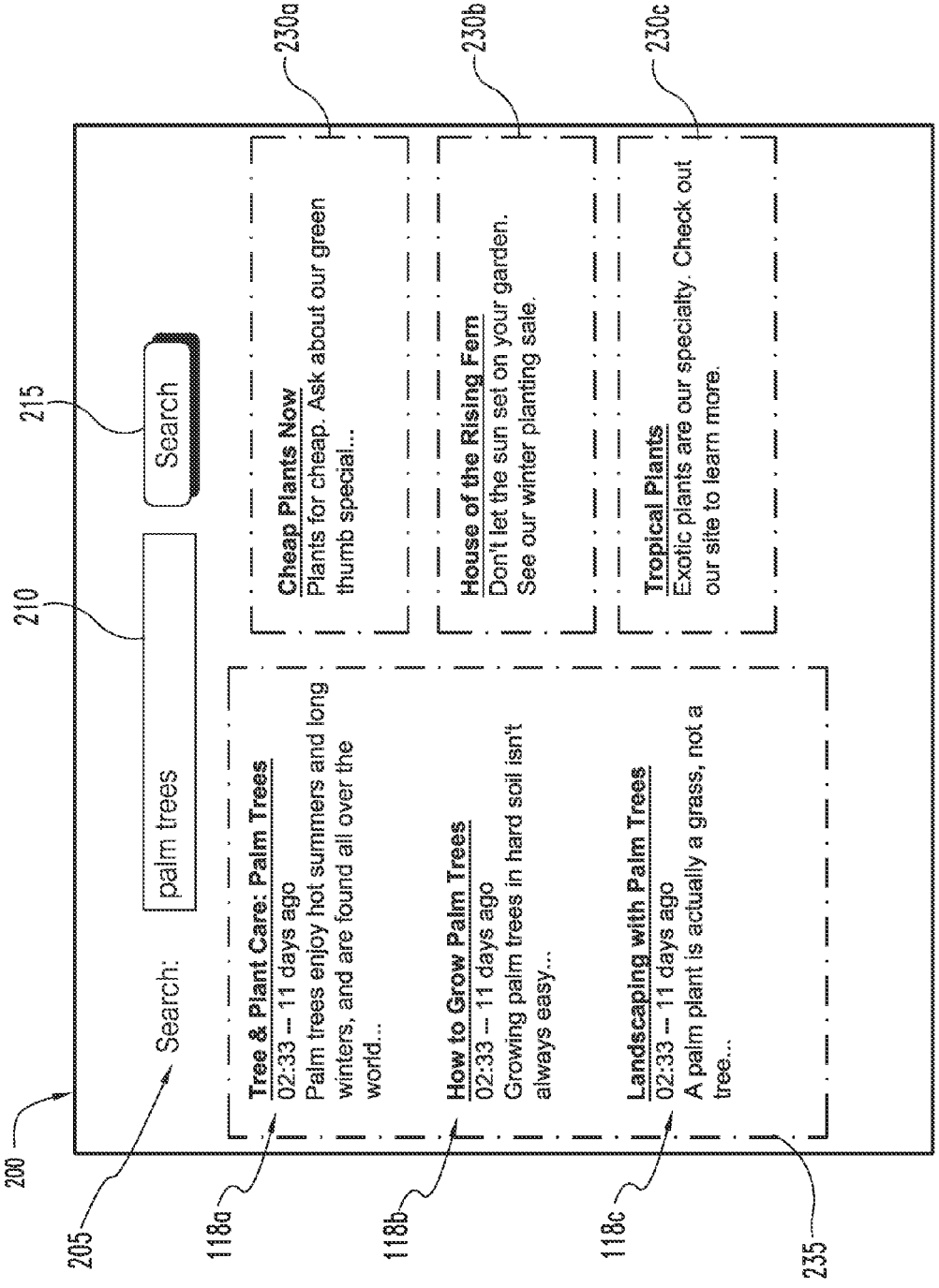

200
Search:
palm trees
210
Search
215
205
Tree & Plant Care: Palm Trees
02:33 -- 11 days ago
Palm trees enjoy hot summers and long winters, and are found all over the world...
118a
How to Grow Palm Trees
02:33 -- 11 days ago
Growing palm trees in hard soil isn't always easy...
118b
Landscaping with Palm Trees
02:33 -- 11 days ago
A palm plant is actually a grass, not a tree...
118c
235
Cheap Plants Now
Plants for cheap. Ask about our green thumb special...
230a
House of the Rising Fern
Don't let the sun set on your garden. See our winter planting sale.
230b
Tropical Plants
Exotic plants are our specialty. Check out our site to learn more.
230c

Fig. 2

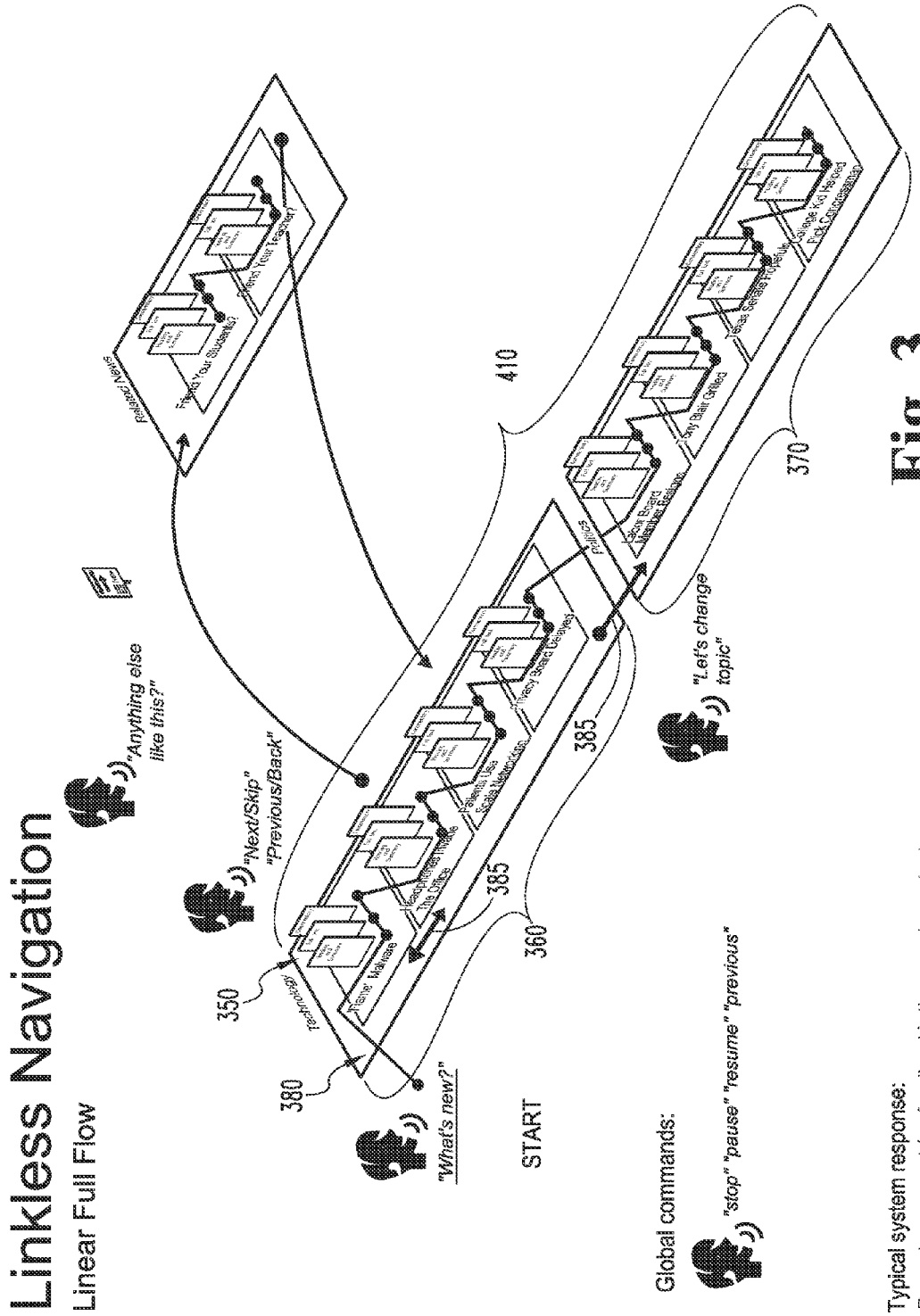

Linkless Navigation
Linear Full Flow
"Anything else like this?"
"Next/Skip"
"Previous/Back"
350
380
"What's new?"
START
360
385
385
410
370
"Let's change topic"
Global commands:
"stop" "pause" "resume" "previous"
Typical system response:
Repeat command (as feedback), then reads content.

Fig. 3

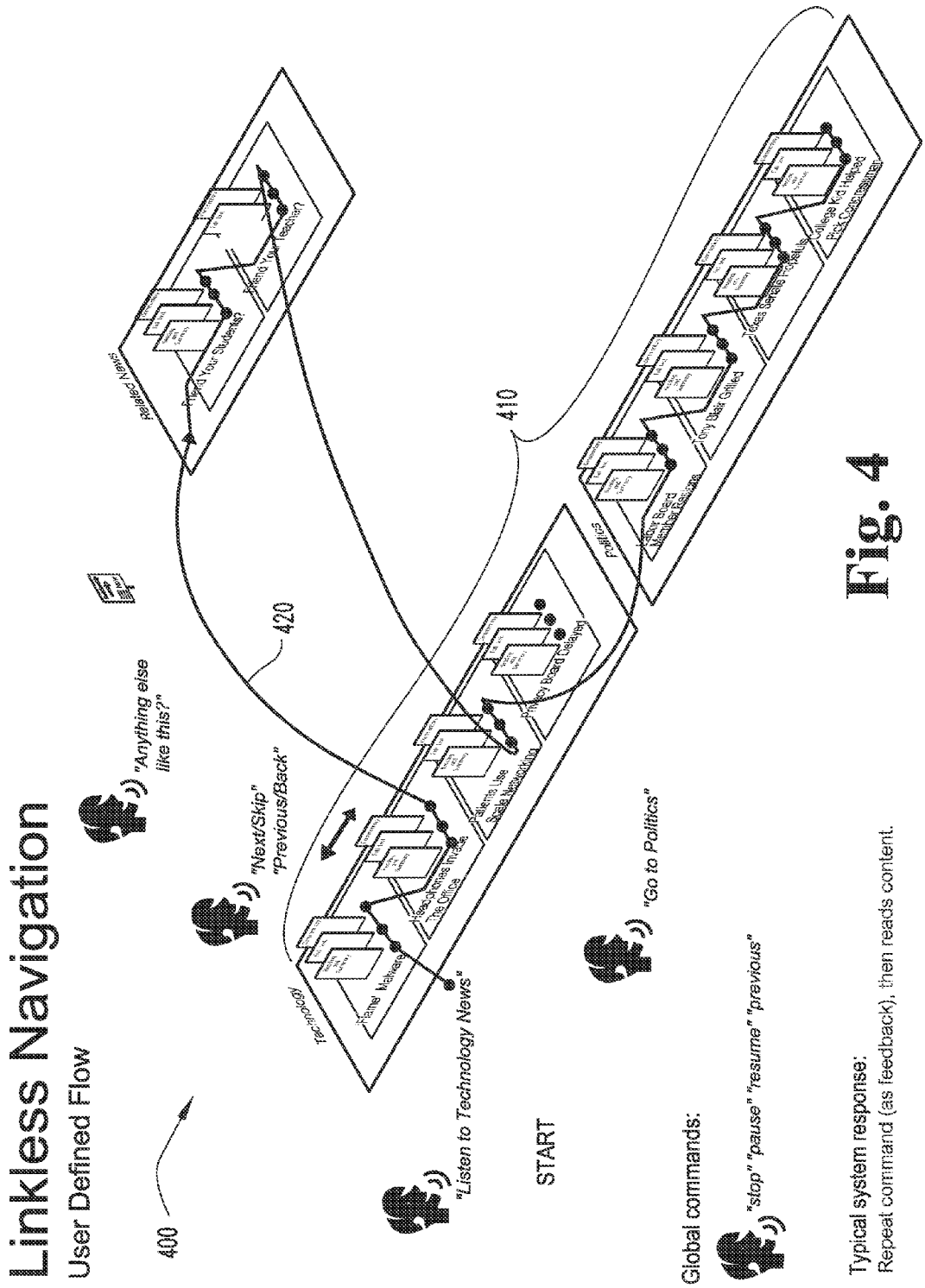

Linkless Navigation
User Defined Flow
400
"Anything else like this?"
"Next/Skip"
"Previous/Back"
420
410
"Listen to Technology News"
START
"Go to Politics"
Global commands:
"stop" "pause" "resume" "previous"
Typical system response:
Repeat command (as feedback), then reads content.

Fig. 4

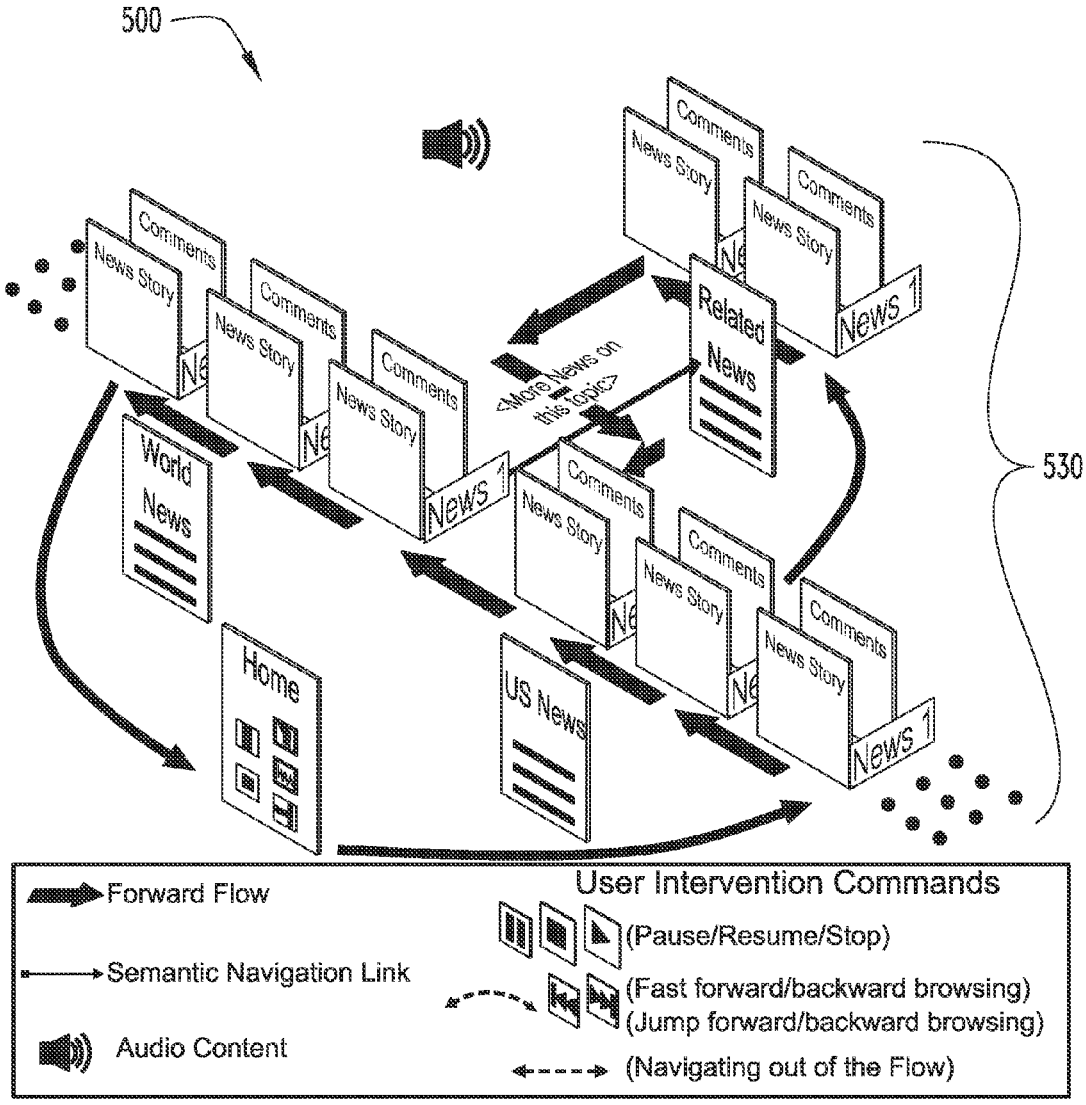

500
530
Comments
News Story
News 1
Related News
<More News on this topic>
World News
Home
US News
Forward Flow
Semantic Navigation Link
Audio Content
User Intervention Commands
(Pause/Resume/Stop)
(Fast forward/backward browsing)
(Jump forward/backward browsing)
(Navigating out of the Flow)

Fig. 5

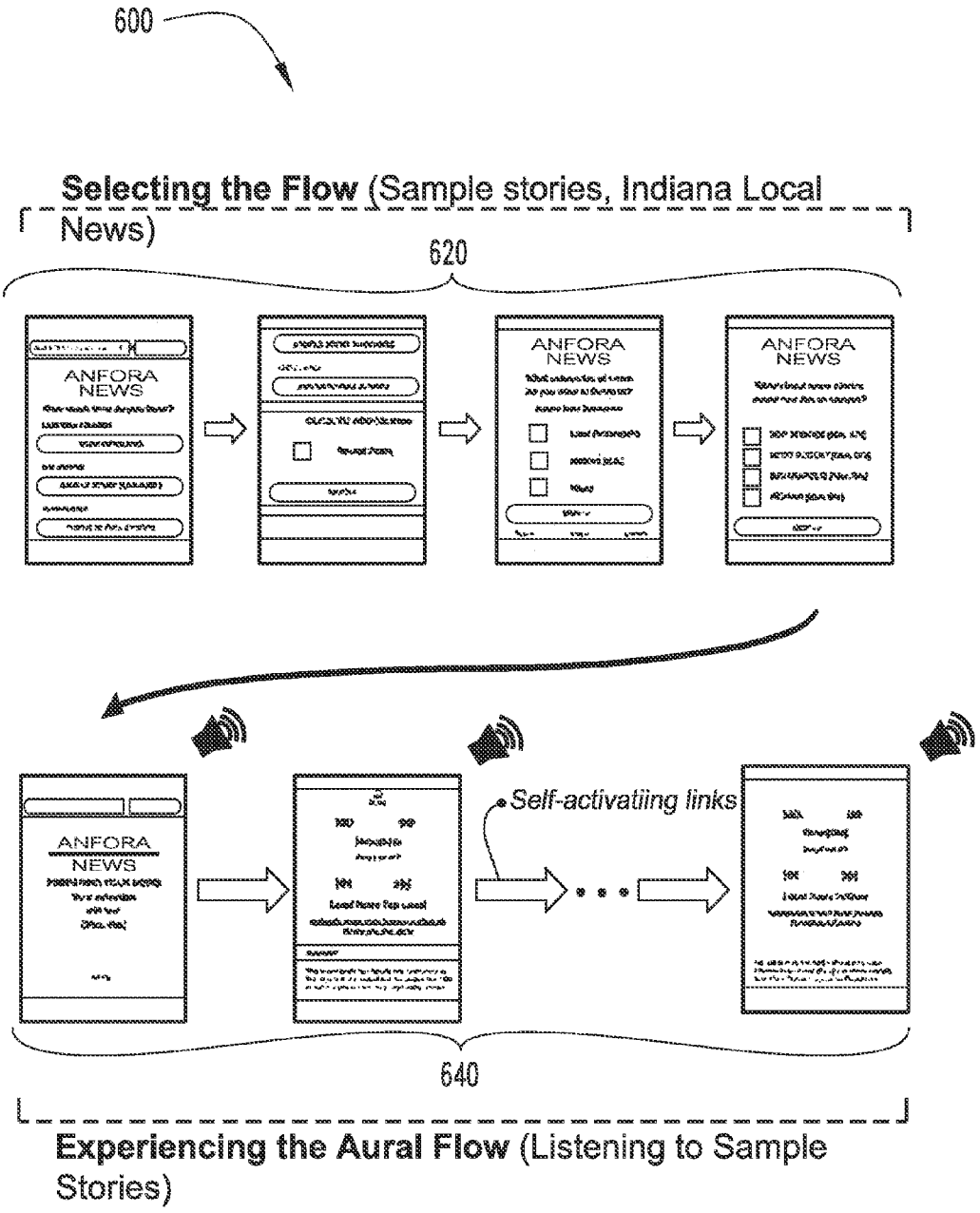
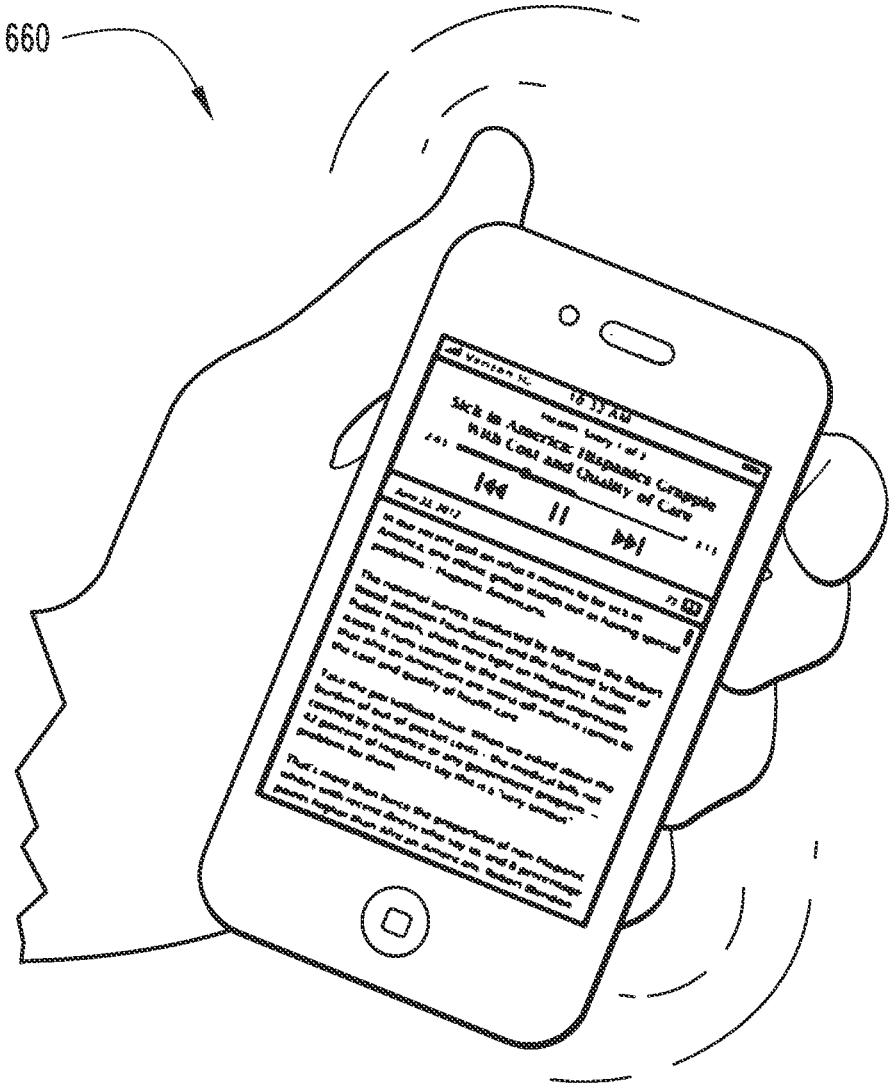

600
Selecting the Flow (Sample stories, Indiana Local News)
620
ANFORA NEWS
ANFORA NEWS
ANFORA NEWS
ANFORA NEWS
Self-activating links
640
Experiencing the Aural Flow (Listening to Sample Stories)

Fig. 6



Accelerometer-based
shake gesture to barge-in
duing aural flow

Fig. 7

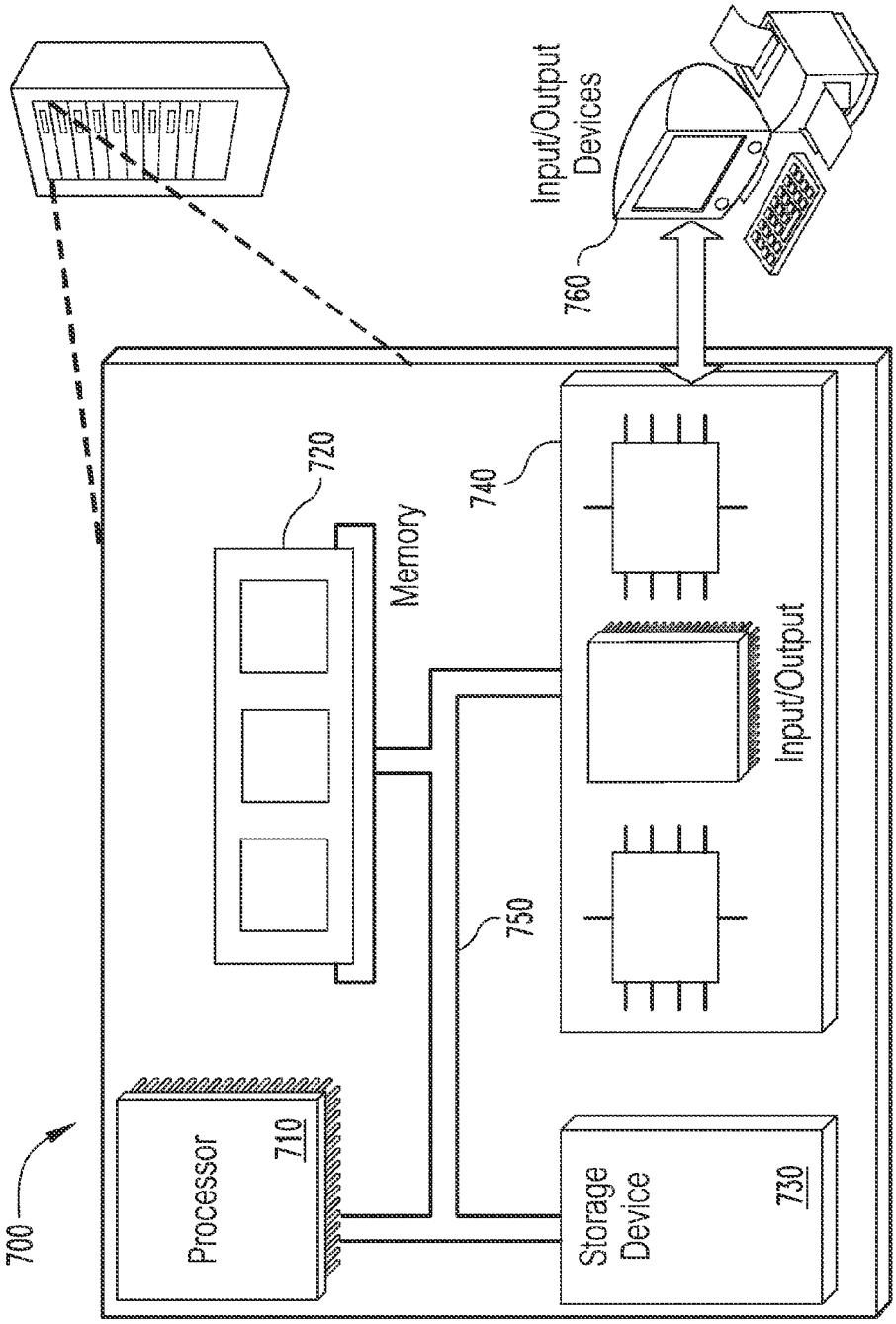

Input/Output Devices
760
720
Memory
740
Input/Output
750
Processor
710
Storage Device
730
700

Fig. 8

AURAL NAVIGATION OF INFORMATION RICH VISUAL INTERFACES

[0001] This patent application claims priority to copending U.S. provisional application No. 61/699,748, filed on Sep. 11, 2012 and incorporates the same herein by reference.

BACKGROUND

[0002] This specification relates to navigation of information and content rich interfaces and applications and specifically the navigation of web based interfaces and applications. Accessing the mobile web on-the-go and in a variety of contexts (e.g., walking, standing, jogging, or driving) is becoming more and more pervasive. Mobile users are often engaged in another activity when it is inconvenient, distracting or even dangerous to continuously look at the web display device at all times. Although existing visual user interfaces can be efficient to support quick scanning of a page, they typically require highly focused attention and may not work well or require a dangerous level of attention in certain situations. It is known that the use of audio-based interfaces of mobile and non-mobile devices during secondary tasks are less distracting and demanding when compared to visual interfaces.

[0003] Another concern is the degree of required or desired interactivity with the web application. Continuous or visually detailed interaction with a conventional web interface requires the user to expend visual attention to the web interface. For example, a user is walking on a city street and would like to catch up with the weekly local news during his 10-minute walk to work. Continuous interaction with a conventional news site on your smart phone would force the user to scan the homepage, ascertain the latest news, selecting a category, potentially followed by selecting a subcategory, and then finally select a news story to read. Once read, the user may want to know more about it or select another news story in the same category, etc. Much of this interactivity is in conflict with the current task of the user's walk to work. Furthermore, the effort expended to both walk and visually interact with the web interface likely amounts to an undesirable user experience. Thus, there is a need for an audio-based system of interaction with data rich interfaces. The present invention addresses this need.

SUMMARY

[0004] This specification describes technologies relating to audio based web navigation and audio web content presentation.

[0005] In general, one innovative aspect of the subject matter described in this specification can be embodied in methods that include the actions of generating a model derived from the analysis of user interactions that represents the hierarchical and hypertextual structures of a website and using that model and user supplied constraints to generate a linear aural flow of content from the said website. An audible presentation based on the linear aural flow is then presented to the user with options for the user to dynamically direct and alter the content of the audio presentation.

[0006] Other embodiments of this aspect include corresponding systems, apparatus, and computer programs, configured to perform the actions of the methods, encoded on computer storage devices.

[0007] The details of one or more embodiments of the subject matter described in this specification are set forth in the accompanying drawings and the description below. Other features, aspects, and advantages of the subject matter will become apparent from the description, the drawings, and the claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0008] FIG. **1** is a block diagram of an example environment in which a paradigm for implementing aural navigation flows on rich architectures manages content delivery services.

[0009] FIG. **2** is an example web page such as might be navigated by an aural navigation system.

[0010] FIG. **3** is a block diagram of an aural navigation system's linear full flow of a collection of web pages.

[0011] FIG. **4** is a block diagram of an aural navigation system's user defined flow of a collection of web pages.

[0012] FIG. **5** is a sample block diagram of a group aural flow in a simplified example web architecture.

[0013] FIG. **6** is a representation of a sample user interface for a mobile device that supports aural navigation flows.

[0014] FIG. **7** is a representation of accelerometer-based shake gesture to interact with an aural flow.

[0015] FIG. **8** is a block diagram of a personal computing device capable of implementing a portion or all of the described technology.

[0016] Like reference numbers and designations in the various drawings indicate like elements.

DETAILED DESCRIPTION

[0017] Before the present methods, implementations and systems are disclosed and described, it is to be understood that this invention is not limited to specific synthetic methods, specific components, implementation, or to particular compositions, and as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular implementations only and is not intended to be limiting.

[0018] As used in the specification and the claims, the singular forms "a," "an" and "the" include plural referents unless the context clearly dictates otherwise. Ranges may be expressed in ways including from "about" one particular value, and/or to "about" another particular value. When such a range is expressed, another implementation may include from the one particular value and/or to the other particular value. Similarly, when values are expressed as approximations, for example by use of the antecedent "about," it will be understood that the particular value forms another implementation. It will be further understood that the endpoints of each of the ranges are significant both in relation to the other endpoint, and independently of the other endpoint.

[0019] "Optional" or "optionally" means that the subsequently described event or circumstance may or may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not. Similarly, "typical" or "typically" means that the subsequently described event or circumstance often though may not occur, and that the description includes instances where said event or circumstance occurs and instances where it does not.

[0020] This application describes a novel, semiinteractive aural paradigm for implementing aural navigation flows on rich architectures enabling users to listen to information-rich interfaces, such as web pages, utilizing complex, hypertextual structures while interacting with the interfaces infrequently. Further, this technology provides for the "aural flow"

and investigates of new ways in which different types of aural flow can be applied to conventional information rich architectures such as web pages. An aural flow is a design-driven, concatenated sequence of pages that can be listened to with minimal interaction required. A flow is governed by aural design rules that determine which pages of the information architecture to automatically concatenate and at which point of the flow the user can interact.

[0021] This technology additionally provides the ability to quickly scanning through content-rich data interfaces, such as web pages, allowing effective but time and/or contextual and/or physical constrained scanning. Finally, the described technology provides a generic design framework applicable to any non-linear, content-rich architecture, such as that which underlies modern web systems. For example, the described technology is appropriate for any large website that features hierarchical and hypertextual structures, such as a commerce, travel planning, or tourism site, and the like.

[0022] FIG. **1** is a block diagram of an example environment **100** in which a paradigm for implementing aural navigation flows on rich architectures manages content delivery services. The example environment **100** includes a network **102**, such as a local area network (LAN), a wide area network (WAN), the Internet, or a combination thereof. The network **102** connects websites **104**, user devices **106** (also known as personal computing device), content sponsors (e.g., advertisers **108**), and an aural navigation system advertisement management system **120**. The example environment **100** may include many thousands of websites **104**, user devices **106** and advertisers **108**.

[0023] A website **104** is one or more resources **105** associated with a domain name and hosted by one or more servers. An example website is a collection of web pages formatted in the hypertext markup language (HTML) that can contain text, images, multimedia content and programming elements, such as scripts. Each website **104** is maintained by a publisher/sponsor, which is an entity that controls, manages and/or owns the website **104**.

[0024] A resource **105** is any data that can be provided over the network **102**. A resource **105** is identified by a resource address that is associated with the resource **105**. Resources include HTML pages, word processing documents, and portable document format (PDF) documents, images, video, and feed sources, to name a few. The resources can include content, such as words, phrases, images and sounds, that may include embedded information (such as meta-information in hyperlinks) and/or embedded instructions (such as JavaScript scripts). Units of content that are presented in (or with) resources are referred to as content items.

[0025] A user device **106** is an electronic device that is under control of a user and is capable of requesting and receiving resources over the network **102**. Example user devices **106** include personal computers, mobile communication devices, and other devices that can send and receive data over the network **102**. A user device **106** typically includes a user application, such as a web browser, to facilitate the sending and receiving of data over the network **102**.

[0026] A user device **106** can request resources **105** from a website **104**. In turn, data representing the resource **105** can be provided to the user device **106** for presentation by the user device **106**. The data representing the resource **105** can also include data specifying a portion of the resource or a portion of a user display (e.g., a presentation location of a pop-up window or in a slot of a web page) in which advertisements can be presented. These specified portions of the resource or user display are referred to as slots or advertisement slots.

[0027] To facilitate searching of these resources **105**, the environment **100** can include a search system **112** that identifies the resources **105** by crawling and indexing the resources **105** provided by the publishers on the websites **104**. Data about the resources can be indexed based on the resource **105** to which the data corresponds. The indexed and, optionally, cached copies of the resources **105** are stored in a search index **114**.

[0028] User devices **106** can submit search queries **116** to the search system **112** over the network **102**. In response, the search system **112** accesses the search index **114** to identify resources that are relevant to the search query **116**. The search system **112** identifies the resources in the form of search results **118** and returns the search results **118** to the user devices **106** in search results pages. A search result **118** is data generated by the search system **112** that identifies a resource that is responsive to a particular search query, and includes a link to the resource. An example search result **118** can include a web page title, a snippet of text or a portion of an image extracted from the web page, and the URL of the web page. Search results pages can also include one or more slots in which other content or advertisements can be presented.

[0029] When a resource **105** or search results **118** are requested by a user device **106**, the advertisement management system **110** receives a request for advertisements to be provided with the resource **105** or search results **118**. The request for advertisements can include characteristics of the slots that are defined for the requested resource or search results page, and can be provided to the advertisement management system **110**.

[0030] For example, a reference (e.g., URL) to the resource for which the slot is defined, a size of the slot, and/or media types that are eligible for presentation in the slot can be provided to the advertisement management system **110**. Similarly, keywords associated with a requested resource ("resource keywords") or a search query **116** for which search results are requested can also be provided to the advertisement management system **110** to facilitate identification of advertisements that are relevant to the resource or search query **116**.

[0031] Based on data included in a given request, the advertisement management system **110** selects advertisements or other content that is eligible to be provided in response to the request (e.g., eligible advertisements). For example, eligible advertisements can include advertisements having characteristics matching those of slots and that are identified as relevant to specified resource keywords or search queries **116**. In some implementations, advertisements that have target keywords that match the resource keywords or the search query **116** are selected as eligible advertisements by the advertisement management system **110**.

[0032] A targeting keyword can match a resource keyword or a search query **116** by having the same textual content ("text") as the resource keyword or search query **116**. The relevance can be based, for example, on root stemming, semantic matching, and topic matching. For instance, an advertisement associated with the targeting keyword "hockey" can be an eligible advertisement for an advertisement request including the resource keyword "hockey." Similarly, the advertisement can be selected as an eligible advertisement for an advertisement request including the search query "hockey."

[0033] A targeting keyword can also match a resource keyword or a search query **116** by having text that is identified as being relevant to a targeting keyword or search query **116** despite having different text than the targeting keyword. For example, an advertisement having the targeting keyword "hockey" may also be selected as an eligible advertisement for an advertisement request including a resource keyword or search query for "sports" because hockey is a type of sport, and therefore, is likely to be relevant to the term "sports."

[0034] The Aural navigation system **120** in some implementations provides a generic design framework applicable to any non-linear, content-rich architecture that is depicted in this example environment **100**. The Aural navigation system **120** provides for aural flows that are modeled on top of existing web information and navigation architectures and can co-exist with the traditional navigation and search mechanisms such as depicted in this example environment **100**. In some implementations, the aural navigation system **120** takes the existing structures and linearizes them appropriately for the aural experience, eliminating the need for changes to the existing websites. For example, the aural navigation system **120** can analyze an existing website **104** such as a news website, and linearize the website for audio presentation such that only simple commands are needed by the user to navigate the audio presentation of the content of the news website.

[0035] In some implementations, the aural navigation system **120** can also utilize user directives, past user browsing and audio browsing history, user stated preferences, and other user information such as user location, user online socio presence, and user schedule when linearizing a website for audio presentation. User directives can be thought of as user supplied defaults. For example, for sites that employ popularity ordering of the article, the user can add defaults to instruct the aural navigation system **120** to ignore articles below a certain ranking. As another example, the aural navigation system **120** can analyze a user's past browsing history to determine that the user typically doesn't review sports articles. Using such information, the aural navigation system **120** could neglect the sports content of a news website **104** when linearizing its content for audio presentation to that user. However, the aural navigation system **120** could override the user's past browsing habits upon encountering sports content that has a significant socio connection with the user. One example of a significant socio connection with the user is the sports content referencing a friend of the user.

[0036] In some implementations, the Aural navigation system **120** is able to perceive and respond to user input (oral or otherwise) and such user input is interpreted within the context of the user's session and user's history. Example commands can include "Change to", "Switch to", "Back", or "Previous" which are sensitive to the users' flow history, not a default flow. Most implementations include various forms of bookmarking enabling the continuing of a story or a topic from a previous session. In some implementations, multiple bookmarks can be maintained enabling the user to go back and continue any of several paused stories. In some implementations, the aural navigation system **120** implements a time-based relevance decay enabling past bookmarked articles to eventually lose their bookmark if not referenced after a period of time.

[0037] Other sample commands include but are not limited to: "What's new?", "Anything else (like this)?", "Next" or "Skip", "Stop" or "Pause", "Resume" "Continue" or "Play" "Listen to" "Go to" "Switch to" or "Change to", "More" or "Tell me more", and "Restart" or "Start over". Note that in some implementations, the aural navigation system **120** is implemented by a user device **106**.

[0038] FIG. **2** is an example web page **200** such as might be navigated by an aural navigation system **120**. The example web page **200** is the resource **105**. The example web page **200** includes a title **205**, a search text slot **210**, a search button **215**, a search results container **235** and advertisement slots **230***a*-**230***c*. The search results container **235** contains the search results **118** of a search performed on this resource. In some implementations, the aural navigation system **120** would provide a "linear flow" of the content of web page **200**, contemplating pre-designated page exits while other implementations provide a "user defined flow" enabling user designated exits and/or content expansion.

[0039] FIG. **3** is a block diagram **300** of an aural navigation system's **120** linear full flow **310** of a collection of web pages **104**. In some implementations of a linear full flow **310**, the flow of information in is strictly linear. Users are able to leave the flow **320** for related stories **330**; upon finishing related stories, they are returned **340** to the original flow. They are only able to jump forward and backward. The flow begins with the first story **350** in the first group of topics **360**. Headline, summary and full story are read in that order. Upon finishing the first story, the system will move on to the next story in that group of topics **360**. Upon finishing the last story in a group of topics **360**, the system will move on to the next group of topics **370**.

[0040] In the block diagram **300**, the lines **380** show the default flow. The lines **320**, **340**, and **385** represent where users can interrupt the flow and move to different parts. The system begins with an orientation cue letting the user know which group of topics they are listening to and the position of the current story in the flow (e.g. "World News, Story 1 of 3). As shown, each story contains a headline, summary, full story and optional related stories.

[0041] In some implementations, the aural navigation system **120** can review the user's browsing history, audio browsing history, location, device **106** usage, socio presence, and calendar when generating a linear full flow of a collection of web pages. For example, a user's browsing history may demonstrate a preference for only the top ranked stories from a particular website **104**. As such, the aural navigation system **120** can anticipate the user's continued browsing pattern by generating a linear flow of the webpages corresponding to the user's anticipated preferences. As another example, a user's browsing pattern could be based upon his location. For example, the content that the user wishes to review in the car can be vastly different than the content that the user wants to review when at work.

[0042] FIG. **4** is a block diagram **400** of an aural navigation system's **120** user defined flow **410** of a collection of web pages. In some implementations, the aural navigation system **120** pauses after reading, audibly disclosing, each dialogue (e.g. summary, full story, reader comments) allowing a user to speak a command. Users can interrupt this flow at any time with any command from the vocabulary. In some implementations, users are able to speak the name of a group of topics (e.g. Politics, U.S., World) and begin the flow in that group. As such, in some implementations each category of content from the website being accessed is available as a command. Categories act as keywords to allowing users the freedom to define their own navigation strategy.

[0043] In this example **400**, line **420** indicates a scenario in which a user leaves the flow to listen to related stories and then changes categories during the flow. Each story contains a headline, summary, full story, reader comments, and two related stories. Users are free to navigate the topics as they please.

[0044] FIG. **5** is a sample block diagram **500** of a group aural flow **530** in a simplified example web architecture. Even in this simplified example, the non-linearity typical nature of such information sources is clearly visible. For example, the example contains different organizational structures (e.g., hierarchical and hypertextual).

[0045] In some implementations, the features of the architecture along with the hypertextual connections are modeled through a collection of primitives and notions known in the art as Interactive Dialogue Module (IDM). IDM provides basic concepts to describe and model hypertextual non-linear architectures. IDM is based on the notion that user interaction can be considered a dialogue between the user and the system. In a nutshell, core content entities (e.g., the news) are multiple topics. A multiple topic can be structured in dialogue acts (news story, commentary on the news story) corresponding to different pages or interaction units composing the topic. Multiple topics are typically organized in groups of topics (e.g., U.S. news or world news) at different hierarchical levels. Hypertextual or semantic associations are typed and can be characterized as structural relationships between multiple topics.

[0046] Using IDM, one or more aural flows are modeled on top of existing web information and navigation architectures as represented by IDM. Thus, the aural flows can co-exist with the traditional navigation and interaction paradigm. As a more complete explanation, an aural flow can be thought of as a design-driven, concatenated sequence of web pages that can be listened to with minimal interaction. The flow is governed by aural design rules that determine which pages of the information architecture to automatically concatenate and at which point of the flow the user can interact. Such design rules can be proposed and refined through various machine learning statistical techniques. For example, concatenation rules can be derived from topic popularity as determined by related topic page hits. Or as another example, statistical models can be derived from topic popularity measures and web activity measures. Similar to predicting conversion for a sales event, the popularity measures and activity measures can be used to derive a "conversion-like" predictor capable of providing a predictive expectation value for topic popularity.

[0047] In some implementations, the user is presented with two flow patterns. The user may either follow the Default Full Flow with little to no interaction, or they may navigate where they please within the flow, creating their own User-Defined Flow. The Default Full Flow, unless interrupted by the user, follows a linear, concatenated flow of information. Typical implementations provide the headline and summary of articles and then provide a portion of the content based upon the aural flow rules. For example, the aural flow rules could provide the full content along with the commentary. The flow continues for each content or story deemed to be above a certain threshold of interest or automatically included by a default behavior. Upon finishing a content, the system will move on to the next content in that group of topics. Upon finishing the last story in a group of topics, the system will move on to the next group of topics. The next group of topics can be based upon the underlying web architecture, derived interest rules (where a topic may have a perceived higher interest than another), or from user derived interest rules (such rules can be derived from previous user actions or directly obtained through a user initiation where the user acts provide rules to govern topic interest).

[0048] However, a user can interrupt the default flow at any time with command from the vocabulary (e.g. "stop" or "change to"). They may navigate wherever they please, at any time they want. This freedom of control creates a User-Defined Flow. An important feature of this flow type is that the system will keep track of a user's history and context during each session. For example, saying a command like "Previous" will take them to the last story they heard, not the previous story in the default flow. In some implementations, the User-Defined Flow still follows the order of the Default Full Flow until a user utters a command. A table of example commands and their respective actions are presented below. In most implementations, users have at least four basic categories of interaction. The four categories are a) Pause, resume, replay and stop: The user can pause and resume the flow. The same dialogue act can be replayed from the beginning. The user can also stop the flow to go back to the home page, b) Fast forward/backward browsing: The user can fast forward to go to the next dialogue act of the same topic or fast backward to go back to the previous dialogue act of the same topic, c) Jump forward/backward browsing: The user can jump forward to the next topic or jump backward to the previous one at any time, d) Navigating out of the flow: The user might want to listen to the related topic by clicking on its link. This action breaks the current flow and moves outside the flow to the desired content (e.g., Related News).

[0049] Note that some implementations provide for a preliminary input or presentation guiding input from a user. Table 2 provides an example of the different characteristics of aural flow types. This preliminary input enables the system to tailor the aural flow to the user's current expectations and/or limitations. For example, in some implementations the user can tell the system an amount of time that the user has with which to listen. For example, the user can tell the system that he has 20 minutes, in which the aural flow through possible content will be streamlined. As another example, a user after choosing a main group of topic, such as U.S. news, could listen to all of the headlines or story summaries in that category. Users would be able to navigate through all the news stories in one category and continue the flow with the next category of news or related stories.

[0050] It has been observed that two sources of error account for a sizable portion of the errors between user and aural flow interaction. The two sources of errors are speech recognition errors and navigation errors. Recognition errors occur when the system either does not understand the uttered command, or the uttered command is not within the scope of the command vocabulary. In some implementations, recognition errors are handled through the notification of the user by the system emitting an earcon, a distinct and noticeable. After being such notified, the user can then reissue the command or issue a different command.

[0051] It is worth mentioning that some implementations provide for a hybrid interface consisting of both the audio presentation along with a visual interface dynamically cued to the content of the current flow. Such hybrid interfaces enable the "At-a-glance" visual confirmation of content. Additionally, such implementations provide for a more extensive visual coverage of the current topic. Such implementa-

tions typically provide an interactive mechanism, for example, a swipe on the personal computing device's touch sensitive screen, to visually provide the full coverage of the topic currently being disclosed.

[0052] Navigation errors occur when a user utters a command that is not applicable in the current part of the flow (e.g. saying "Forward" while in the commentary"). These errors should be handled with audio orientation cues provided by the system (e.g. "There is no more content for this story"). In some implementations, the system responds by reverting back to a default flow. Alternatively, some implementations respond by audibly providing a shortened menu of possible actions based upon the user's current location in the flow.

[0053] FIG. **6** is a representation of a sample user interface **600** for a mobile device that supports aural navigation flows. The aural flow experience consists of two main components, as highlighted in the figure: Selecting the flow **620** and experiencing the flow **640**. In Selecting the flow **620**, the system provides to the user several options to choose and customize the coverage of the available content, based on time constraints, types of aural flow and user's interest. A simple sequence of user interface screens is shown supporting this selection task. Once the user has selected values for such simple parameters, the system immediately generates and makes available the aural flow corresponding to the user's selection. At this point, the user enters the Experiencing the flow **640** part, in which the system plays the aural flow, which concatenates the web pages through self-activating links.

[0054] FIG. **7** is a representation **660** of accelerometer-based shake gesture to interact with an aural flow. In some implementations, the aural flow can be interacted and altered through vocal and/or tactile user actions and/or a locational value of the personal computing device. In such implementations, activating a microphone will temporarily stop the system output and activate the "listening mode." During this pause, the system will wait for a command. If the button is released with no command having been uttered, the system will simply resume its output. If a command was uttered and understood by the system, the system will react accordingly. Shaking the personal computing device and utilizing the accelerometer to activate the listening mode works similarly to curing the microphone.

[0055] The locational input utilizes a geographical positional system component that is typical of many personal computing devices. However, the locational input functionality differs from the input of direct user actions. Locational input is typically configured by the user to respond in certain ways to locational values. For example, a user could configure the system such that the content, as presented upon arriving at the user's place of employment, consist of the latest topics on the company's intranet. As another example, a user could request that the content be based upon the user's geographic position.

[0056] FIG. **8** is a block diagram of a personal computing device **700** capable of implementing a portion or all of the described technology. The example of one such type of personal computing device **700** shows a block diagram of a programmable processing system (system) **700** suitable for implementing apparatus or performing methods of various aspects of the subject matter described in this specification. The system **700** includes a processor **710**, a random access memory (RAM) **721**, a program memory **730** (for example, a writable read-only memory (ROM) such as a flash ROM) and an input/output (I/O) controller **740** (typically endowed with GPS capability) coupled with a bus **750**. The system **700** can be preprogrammed, in ROM, for example, or it can be programmed (and reprogrammed) by loading a program from another source (for example, downloaded from an application site, or another personal computing device).

[0057] The I/O controller **740** is operationally connected to I/O interfaces **760**. The I/O interface receives and transmits data (e.g., stills, pictures, movies, and animations for importing into a composition) in analog or digital form over communication links such as a serial link, local area network, wireless link, and parallel link, cellular, touch and shake inputs, geographic locational input, and the like.

TABLE 1

Sample system navigation and commands (primitives)

Command	System	System Action
What's new?"	"Recent stories in (topic) news"	Begin default full flow
"Anything else (like this)?" "More like this"	"Related stories"	Go to related stories
"Next" or "Skip"	"Next story"	Go to next story
"Previous" or "Back"	"Previous story"	Go to previous story in user history
"Stop" or "Pause"	Earcon	Pauses story
"Resume" "Continue" or "Play"	"Resuming (headline)"	Resumes story
"Listen to" "Go to" "Switch to" or "Change to"	"Switching to (topic) news"	Switch to selected topic
"Forward" or "Rewind"	Title of next section is read	Move between sections within a story
"Restart" or "Start over"	"Restarting (reads headline)"	Restarts story

TABLE 2

example of the different characteristics of aural flow types

Flow	Characteristics	Time	Advantages	Disadvantages
Group	A selected group of topics	5 min	Decide the category from the outset	Interact every time to select a different category
Full	All groups of topics	Longer period of time - 30 min.	Less interaction	Difficulty building mental model
Deep	All groups of topics + semantic associations	Longer period of time - 1 hr.	In-depth coverage of content	Difficulty building mental model
Light	Agile overview of each topic (default dialogue act)	Shorter period of time	More stories in less time (agile overview)	Details of each topic will not be played
Rich	Extensive coverage of each topic (all dialogue acts)	Longer period of time	Extensive coverage	Time-consuming and constraining

[0058] Embodiments of the subject matter and the operations described in this specification can be implemented as a method, in digital electronic circuitry, or in computer software, firmware, or hardware, including the structures dis-

closed in this specification and their structural equivalents, or in combinations of one or more of them. Embodiments of the subject matter described in this specification can be implemented as one or more computer programs, i.e., one or more modules of computer program instructions, encoded on computer storage medium for execution by, or to control the operation of, data processing apparatus. Alternatively or in addition, the program instructions can be encoded on an artificially-generated propagated signal, e.g., a machine-generated electrical, optical, or electromagnetic signal that is generated to encode information for transmission to suitable receiver apparatus for execution by a data processing apparatus. A computer storage medium can be, or be included in, a computer-readable storage device, a computer-readable storage substrate, a random or serial access memory array or device, or a combination of one or more of them. Moreover, while a computer storage medium is not a propagated signal, a computer storage medium can be a source or destination of computer program instructions encoded in an artificially-generated propagated signal. The computer storage medium can also be, or be included in, one or more separate physical components or media (e.g., multiple CDs, disks, or other storage devices).

[0059] The operations described in this specification can be implemented as operations performed by a data processing apparatus on data stored on one or more computer-readable storage devices or received from other sources.

[0060] The term "data processing apparatus" encompasses all kinds of apparatus, devices, and machines for processing data, including by way of example a programmable processor, a computer, a system on a chip, or multiple ones, or combinations, of the foregoing The apparatus can include special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application-specific integrated circuit). The apparatus can also include, in addition to hardware, code that creates an execution environment for the computer program in question, e.g., code that constitutes processor firmware, a protocol stack, a database management system, an operating system, a cross-platform runtime environment, a virtual machine, or a combination of one or more of them. The apparatus and execution environment can realize various different computing model infrastructures, such as web services, distributed computing and grid computing infrastructures.

[0061] A computer program (also known as a program, software, software application, script, or code) can be written in any form of programming language, including compiled or interpreted languages, declarative or procedural languages, and it can be deployed in any form, including as a stand-alone program or as a module, component, subroutine, object, or other unit suitable for use in a computing environment. A computer program may, but need not, correspond to a file in a file system. A program can be stored in a portion of a file that holds other programs or data (e.g., one or more scripts stored in a markup language document), in a single file dedicated to the program in question, or in multiple coordinated files (e.g., files that store one or more modules, sub-programs, or portions of code). A computer program can be deployed to be executed on one computer or on multiple computers that are located at one site or distributed across multiple sites and interconnected by a communication network.

[0062] The processes and logic flows described in this specification can be performed by one or more programmable processors executing one or more computer programs to perform actions by operating on input data and generating output. The processes and logic flows can also be performed by, and apparatus can also be implemented as, special purpose logic circuitry, e.g., an FPGA (field programmable gate array) or an ASIC (application-specific integrated circuit).

[0063] Processors suitable for the execution of a computer program include, by way of example, both general and special purpose microprocessors, and any one or more processors of any kind of digital computer. Generally, a processor will receive instructions and data from a read-only memory or a random access memory or both. The essential elements of a computer are a processor for performing actions in accordance with instructions and one or more memory devices for storing instructions and data. Generally, a computer will also include, or be operatively coupled to receive data from or transfer data to, or both, one or more mass storage devices for storing data, e.g., magnetic, magneto-optical disks, or optical disks. However, a computer need not have such devices. Moreover, a computer can be embedded in another device, e.g., a mobile telephone, a personal digital assistant (PDA), a mobile audio or video player, a game console, a Global Positioning System (GPS) receiver, or a portable storage device (e.g., a universal serial bus (USB) flash drive), to name just a few. Devices suitable for storing computer program instructions and data include all forms of non-volatile memory, media and memory devices, including by way of example semiconductor memory devices, e.g., EPROM, EEPROM, and flash memory devices; magnetic disks, e.g., internal hard disks or removable disks; magneto-optical disks; and CD-ROM and DVD-ROM disks. The processor and the memory can be supplemented by, or incorporated in, special purpose logic circuitry.

[0064] To provide for interaction with a user, embodiments of the subject matter described in this specification can be implemented on a computer having a display device, e.g., a CRT (cathode ray tube) or LCD (liquid crystal display) monitor, for displaying information to the user and a keyboard and a pointing device, e.g., a mouse or a trackball, by which the user can provide input to the computer. Other kinds of devices can be used to provide for interaction with a user as well; for example, feedback provided to the user can be any form of sensory feedback, e.g., visual feedback, auditory feedback, or tactile feedback; and input from the user can be received in any form, including acoustic, speech, or tactile input. In addition, a computer can interact with a user by sending documents to and receiving documents from a device that is used by the user; for example, by sending web pages to a web browser on a user's client device in response to requests received from the web browser.

[0065] Embodiments of the subject matter described in this specification can be implemented in a computing system that includes a back-end component, e.g., as a data server, or that includes a middleware component, e.g., an application server, or that includes a front-end component, e.g., a client computer having a graphical user interface or a Web browser through which a user can interact with an implementation of the subject matter described in this specification, or any combination of one or more such back-end, middleware, or front-end components. The components of the system can be interconnected by any form or medium of digital data communication, e.g., a communication network. Examples of communication networks include a local area network

("LAN") and a wide area network ("WAN"), an inter-network (e.g., the Internet), and peer-to-peer networks (e.g., ad hoc peer-to-peer networks).

[0066] The computing system can include clients and servers. A client and server are generally remote from each other and typically interact through a communication network. The relationship of client and server arises by virtue of computer programs running on the respective computers and having a client-server relationship to each other. In some embodiments, a server transmits data (e.g., an HTML page) to a client device (e.g., for purposes of displaying data to and receiving user input from a user interacting with the client device). Data generated at the client device (e.g., a result of the user interaction) can be received from the client device at the server.

[0067] While this specification contains many specific implementation details, these should not be construed as limitations on the scope of any inventions or of what may be claimed, but rather as descriptions of features specific to particular embodiments of particular inventions. Certain features that are described in this specification in the context of separate embodiments can also be implemented in combination in a single embodiment. Conversely, various features that are described in the context of a single embodiment can also be implemented in multiple embodiments separately or in any suitable subcombination. Moreover, although features may be described above as acting in certain combinations and even initially claimed as such, one or more features from a claimed combination can in some cases be excised from the combination, and the claimed combination may be directed to a subcombination or variation of a subcombination.

[0068] Similarly, while operations are depicted in the drawings in a particular order, this should not be understood as requiring that such operations be performed in the particular order shown or in sequential order, or that all illustrated operations be performed, to achieve desirable results. In certain circumstances, multitasking and parallel processing may be advantageous. Moreover, the separation of various system components in the embodiments described above should not be understood as requiring such separation in all embodiments, and it should be understood that the described program components and systems can generally be integrated together in a single software product or packaged into multiple software products.

[0069] Thus, particular embodiments of the subject matter have been described. Other embodiments are within the scope of the following claims. In some cases, the actions recited in the claims can be performed in a different order and still achieve desirable results. In addition, the processes depicted in the accompanying figures do not necessarily require the particular order shown, or sequential order, to achieve desirable results. In certain implementations, multitasking and parallel processing may be advantageous.

What is claimed is:

1. A method comprising:

generating, by a computer, a model of a website using user interaction primitives to represent hierarchical and hypertextual structures of the website;

generating, by the computer, a linear aural flow of content of the website based upon the model and a set of user constraints;

audibly presenting, by the computer, the linear aural flow of the content such that the linear aural flow of content is controlled through the use of user supplied primitives, wherein, the linear aural flow can be turned into a dynamic aural flow based upon the user supplied primitives.

2. The method of claim **1** wherein user supplied primitives comprises a spoken command.

3. The method of claim **1** wherein the linear aural flow is further based on a ranking of current topics based upon each topic's page hits the website has received.

4. The method of claim **2** wherein interrupted audibly presented content is bookmarked such that the bookmark ages over a user stated period and is eliminated upon an ending of a user stated period.

5. The method of claim **1** wherein the set of user constraints is derived from a user's past audio browsing history in conjunction with the device used to perform the past audio browsing.

6. The method of claim **1** wherein the user supplied primitives are interpreted in context of a user's session.

7. The method of claim **1** wherein the linear aural flow sequences individual articles into dialogues for audio presentation including a dialog for an article's headline, a dialog for the article's summary, and a dialog for the article's content.

8. The method of claim **2** wherein a spoken command is a name of a category of content available on the website.

9. The method of claim **1** wherein the set of user constraints is derived from popularity measures of articles present on the website.

10. A computer storage medium encoded with a computer program, the program comprising instructions that when executed by a user device cause the user device to perform operations comprising:

receiving a model of a website, the model representing hierarchical and hypertextual structures of the website, wherein the model uses user interaction primitives to represent the hierarchical and the hypertextual structures of the website;

receiving a set of user derived constraints;

generating a linear aural flow of content of the website based upon the model and a set of user derived constraints;

audibly presenting the linear aural flow of the content;

determining whether a user command indicates a desire for a dynamic aural flow;

upon determining that a user command indicates a desire for a dynamic aural flow, audibly presenting a dynamic aural flow.

11. The method of claim **10**, wherein interrupted audibly presented content is bookmarked such that the bookmark ages over a user stated period and is eliminated upon an ending of a user stated period.

12. The method of claim **10** wherein the set of user constraints is derived in part from a user's past audio browsing history in conjunction with the device used to perform the past audio browsing.

13. A system comprising:

a user device;

one or more computers operable to interact with the device;

instructions stored on a machine readable storage device for execution by the one or more computers, wherein upon execution the instructions cause the one or more computers to perform the operations of:

generate a model of a website, the model representing hierarchical and hypertextual structures of the website through usage of user interaction primitives;

generate a linear aural flow of content of the website based upon the model and a set of user constraints;

provide instructions to the user device causing the user device to audibly present the linear aural flow of the content;

upon receiving input from a user, provide instructions to the user device causing the user device to audibly present a dynamic aural flow of the content.

14. The system of claim **13**, wherein the one or more computers comprise the user device.

15. The system of claim **13**, wherein the linear aural flow is further based on a ranking of current topics based upon each topic's page hits the website has received.

16. The system of claim **13**, wherein the one or more computers comprise a server operable to interact with the device through a data communication network, and the user device is operable to interact with the server as a client.

17. The system of claim **13**, wherein the one or more computers consist of one computer, the user device is a user interface device, and the one computer comprises the user interface device.

* * * * *