Technical Disclosure Commons

Defensive Publications Series

July 03, 2017

Actions on Landing Pages

John D. Lanza
Foley & Lardner LLP

John D. Lanza
Foley & Lardner LLP

ACTIONS ON LANDING PAGES

ABSTRACT

Systems and methods described herein allow for using audio overlay to provide actionable audio extensions with primary audio content. A data processing system can receive a first input audio signal from a client device of a user, and identify a user request from the first input audio signal. The data processing system can generate an audio response to the user's request, and one or more audio extensions for presenting with the audio response. The data processing can then transmit the audio response and the audio extensions to the client device. The data processing system may receive a second audio signal indicative of an interaction with one of the audio extensions. In response, the data processing system can execute an operation associated with the interaction provided in the second audio signal.

DETAILED DESCRIPTION

In a voice activated computer environment, information can be provided to users in audio format. For instance, users can initiate audio conversations with a voice-based virtual personal assistant (VPA) to request services, and the voice-based VPA can respond back with audio data. The audio data can include information related to a requested service, additional information, requests for additional information, requests to confirm certain information, or a combination thereof. Additional information (e.g., URLs of landing pages, phone numbers, e-mail addresses, geographical information, reviews of a product or service, reviews of a product or service provider,) and/or sponsored content (e.g., ads) can be provided with primary content (e.g., information requested by a user) for presenting to users.

The present disclosure aims to improve user experience and increase user engagement with audio content presented during conversations between users and the voice-based VPA. In particular, a data processing system can generate actionable audio extensions to primary content and/or sponsored content, and provide the generated audio extensions for presenting to users during conversations with the voice-based VPA. The actionable audio extensions can facilitate or enhance user interactions or engagements with presented content. For example, the actionable audio extensions can be indicative of phone numbers, email addresses, URLs, reviews, or geographical data related to requested services or sponsored content, and users can select to use such audio extensions to execute transactions or access additional information.

When presented as visual content, users can easily distinguish actionable extensions from primary content or other sponsored content. For instance, some actionable item may be presented with a different visual format (e.g., different color, underlined, different font, etc.) than other content presented to the user. For some actionable extensions, when a user hovers a computer device cursor on an actionable item, the cursor's shape may change to indicate that the item is actionable (e.g., clickable). An actionable item may also include text (e.g., "click to call" or "swipe for more details") to instruct users to interact with it. However, such visual formatting options may not be adequate for audio extensions. Furthermore, user interaction may take another form when content is presented in audio format.

Methods and systems described herein can include a data processing system that can identify a creative for serving as audio content to a user, responsive to an identified user request. The data processing system may receive the creative from a service/product provider or an advertiser, or retrieve such extension from a website or database associated with a service/product provider. The data processing system can transform a visual extension to a

corresponding audio extension and/or apply some sound or audio effects to distinguish the generated audio extension from other audio content presented to the user. The data processing system may provide the audio extension for presentation as an audio overlay. The audio extension may include audio content for asking the user whether to actuate the extension. The data processing system may monitor user responses regarding interacting with audio extensions. The data processing system may monitor user interactions across multiple devices. The data processing system may record user interaction events or report them to other computing devices.

FIG. 1 is a flowchart depicting an example method 100 for providing actionable audio extensions in a voice activated computer network environment. The method 100 can include, at step 105, identifying a user request based on a received input audio signal. A user can initiate, via a respective client device, a conversation with an instance of a voice-based VPA, and utter a request or command for a service or online action (e.g., an audio search query, request for making or checking an airline flight reservation, request for online purchase of movie tickets, request for rendering of an online live stream, request to schedule a cab service, etc.). The client device can receive the input audio signal via a respective microphone, and transmit the received input audio signal to the data processing system. A natural language processor (NLP) component of the data processing system can process the input audio signal to identify the user request. The NLP component can machine-translate the input audio signal to a corresponding text and parse the generated text to identify one or more keywords.

For example, the audio signal detected by the client device can include “OK, I would like to go to go dinner.” In such example, the data processing system can identify the keywords “I need,” “go to” and/or “dinner,” and determine a user request for a restaurant to have dinner. The data processing system can interpret the user request as a request for suggestions regarding

restaurants within a certain distance from the user's location. In another example, the input audio signal can include "I want to watch the Team 1 – Team 2 game." The data processing system can identify the keywords "watch," "game," "Team 1," and "Team 2," and determine that the user is requesting a streaming service to watch the soccer game between Team 1 and Team 2.

At step 110, the data processing system can generate an audio response to the identified user request. For example, the data processing system can perform an online search for restaurants within a geographical area associated with the user. The data processing system may acquire from the user client device a respective geographical location. The data processing system may also access a user profile to determine user preferences with respect to restaurants. The data processing system may generate a search query using keywords indicative of a geographical location and/or one or more types of restaurants, and transmit the search query to a search engine or a website related to restaurants. In response to the transmission of the search query, the data processing system can receive search results indicative of a list of restaurants. Each item in the search results can include a restaurant name, a brief description, a URL of a corresponding landing page, and/or a geographical address. In another example, the data processing system can generate a search query to search for streaming sessions for the soccer game requested by the user, and provide the search query to search engine and/or computing devices of one or more streaming service providers. The data processing can receive search results including indications of the streaming service providers, qualities of streamed content, and/or prices.

The data processing system can also select one or more ads for presenting with the search results. For example, the data processing system may select an ad for a restaurant not within the received search results (e.g., a restaurant that is not within the geographical location associated

with the user or does not conform to the user preferences). The data processing system may select, in another example, an ad for a pay-per-view service related to the soccer game requested by the user. In general, the ad may be related to an alternative service provider (e.g., other than providers listed in the search results or other than a service provider specified by the user), an enhanced version of a requested service, or other products/services related to the user requested service.

The data processing system can generate the audio response based on search results and/or other information received from computing devices of one or more service providers. The data processing system may employ an audio generating component to transform text content (e.g., including names and/or descriptions) associated with the search results, the selected ad, and/or information received from a service provider to corresponding audio content. The generated response can include multiple audio items. For example an audio item in the response can be indicative of a corresponding item in the search results or the selected ad.

At step 115, the data processing system can generate an audio extension for presenting with the generated audio response. The data processing system can access a URL associated with a search results item, an ad, and/or a specific service provider, and analyze content of the corresponding webpage. The data processing system can extract information or sitelinks from the corresponding webpage. For example, the data processing system can retrieve from a restaurant website information indicative of a phone number, a geographical address, and/or an email address associated with the restaurant. The data processing system may also identify one or more links for content indicative, for example, of the restaurant menu, reviews, and/or driving directions. In some implementations, the data processing system may retrieve the information for generating the extension from an ad campaign database.

In generating the audio extension, the data processing system may not necessarily include all or the exact information extracted from the landing page. For instance, the data processing system may generate audio labels indicative of different pieces of information obtained from the landing page or the ad campaign. In particular, the data processing system may generate audio segments corresponding to the labels "call," "email," "menu," and/or "driving directions," and append the generated response with the audio segments corresponding to these labels (or a combination thereof). The data processing system may also create a data structure that maps each of the labels to respective information and/or actions. For example, the label "call" may be mapped to a phone number of the service provider and one or more actions/operations including (i) calling the phone number and (ii) providing the phone number on a client device. The label "menu" may be mapped to a corresponding link and one or more actions/operations including (i) generating and providing an audio version of the menu and (ii) providing a visual version of the menu for display on a client device.

The data processing system may apply sound or audio effects, such as a flanging effect, an equalization effect, audio filters, a compression effect, a robotic voice effect, pitch manipulation, or a combination thereof, to the audio segments corresponding to the labels to distinguish (audibly) the labels from other audio content in the generated response. The data processing system may also (or alternatively) use background audio and/or tones before or after the audio segments corresponding to the labels. The data processing system may generate multiple audio extensions with such labels. For example, the data processing system can create an audio extension for each content item (e.g., search result item, service provider item, or ad) in the generated response. In such example, each content item of the generated may be appended

with a corresponding audio extension (or audio labels). The audio extensions can be viewed as an audio overlay to the audio response.

At step 120, the data processing system can provide the generated audio response and the audio extension to the client device for presenting to the user. The data processing system can transmit the generated audio response and the audio extension as one or more audio signals to the client device. When the client device plays the audio extension (or audio labels) associated with a given content item (e.g., search result item, ad, or service provider item) of the audio response, the user of the client device can utter one of the labels as an indication of an interaction intent. For example, the user may the label "call." The user may also specify an action/operation associated with the label. For example, the user may utter the words "display menu on my phone." The client device can transmit the audio signal corresponding to the user's utterance to the data processing system. The data processing system can machine-translate the received audio signal to a corresponding text signal, and use the generated text signal to map the user's utterance to one of the labels and/or a corresponding action/operation in the data structure. Responsive to the mapping, the data processing can the execute the action/operation requested by the user (e.g., cause a user device to call the service provider's phone number, display the phone number on a user's device, provide the link corresponding to the menu for display on a user device, provide the menu in audio format to the client device, etc.). The data processing system may provide content to be displayed on a user device different than the client device from which audio signals were received.

The data processing system may record an indication of user interaction with the audio extension responsive to mapping the user's utterance to a corresponding label and/or a respective action/operation. The data processing system may transmit such an indication to a computing

device associated with an advertiser or a service provider. The data processing system may cause statistical historical data corresponding to the user (e.g., related to user interactions) or to a service provider to be updated responsive to mapping the user's utterance to a corresponding label and/or a respective action/operation. The data processing system may also receive from a user device indications of user's interactions with visual content resented responsive to mapping the user's utterance to a corresponding label and/or a respective action/operation. The data processing system can use collected data indicative of user interactions to update user preferences and/or to select content for presenting to the user in future conversations.

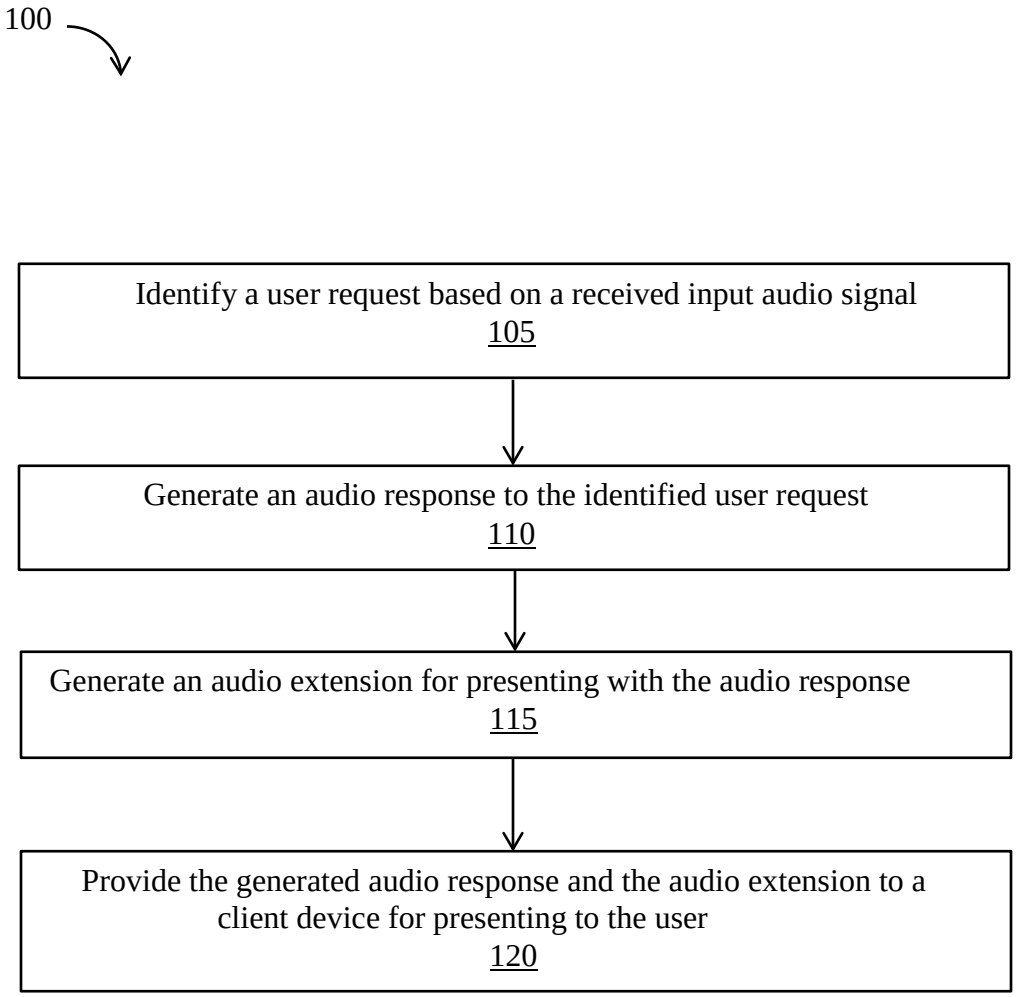
100
Identify a user request based on a received input audio signal
105
Generate an audio response to the identified user request
110
Generate an audio extension for presenting with the audio response
115
Provide the generated audio response and the audio extension to a client device for presenting to the user
120

FIG. 1