Technical Disclosure Commons

Defensive Publications Series

October 2020

Real-time Contextual Searches to Assist Speakers

Dheeraj Agrawal

Piyush Itankar

Real-time Contextual Searches to Assist Speakers

ABSTRACT

Persons who speak professionally to audiences, e.g., teachers, typically spend substantial time and effort to prepare presentation or lecture materials. Answers to follow-up questions from the audience can become more meaningful if the speaker has access to relevant backup material, which, however, may not always be on hand. This disclosure describes techniques to assist speakers in a dynamic, real-time, and automatic manner. Speaker presentation materials and spoken statements are obtained with permission and are used to formulate contextual queries. Results of the queries that include contextually relevant content are grouped by topic and are displayed in a user interface, e.g., as a sidebar that is viewable by the speaker. The results can be displayed in a tabbed user interface that separates content based on content type.

KEYWORDS

- Real-time search
- Speech-triggered search
- Speech-to-text
- Lookahead search
- Contextual search
- Contextual AI
- Artificial intelligence (AI)

BACKGROUND

Persons who speak professionally to audiences, e.g., teachers, typically spend substantial time and effort to prepare presentation or lecture materials. They strive to overcome the challenge of conveying ideas that their students or audiences may be hearing or seeing for the first time. Answers to follow-up questions from the audience can become more meaningful if the speaker has access to relevant backup material, which, however, may not always be on hand.

DESCRIPTION

This disclosure describes techniques to assist speakers, e.g., teachers, etc., in a dynamic, real-time, and automatic manner by gathering and displaying, e.g., as a side-bar viewable by the speaker, information on the subject of the presentation. Information can be gathered generally from the Internet, and particularly from sessions similar to the one being delivered by the speaker when recordings or materials such sessions are available. The techniques leverage widely available resources to enable speakers to rapidly develop a better representation of their topics in the form of slides, documents, articles, videos, images, etc. Speakers can use this information, derived from live AI, to discuss, teach, resolve audience or students' queries, etc.

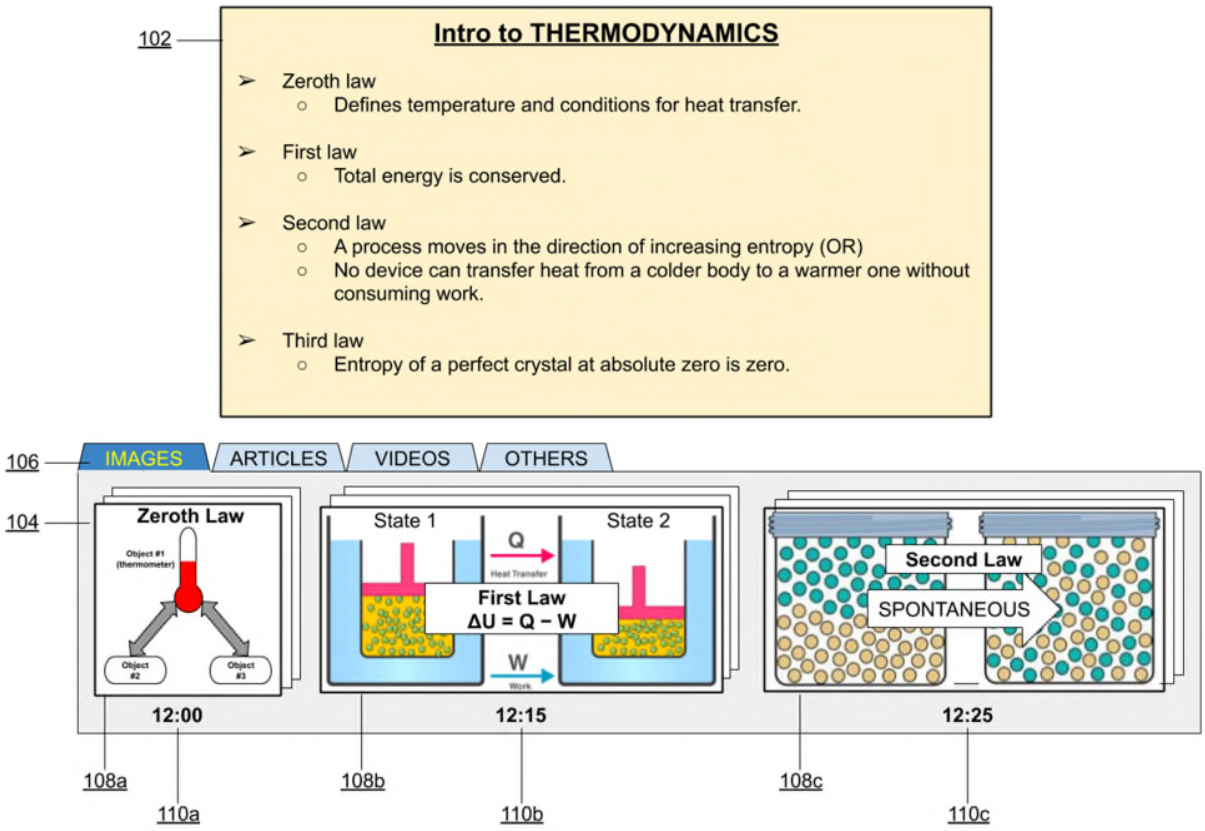

Fig. 1: Real-time, automatic, contextual searches to assist teachers, speakers, or presenters

Fig. 1 illustrates real-time, automatic, contextual searches to assist teachers, speakers, or other presenters, per the techniques of this disclosure. In presentation mode, the speaker displays to an audience a document or a slide (102) from a set of slides. The speaker's view of the presentation includes results (108a-c) in a moving bottom bar (104) (or sidebar) obtained from automatic, real-time searches relating to the content of the slides.

For example, if the topic of the slide is "thermodynamics," (as shown in Fig. 1) with bullet points that include the zeroth, first, second, and third laws of thermodynamics, then the searches pull up and display on the moving sidebar material, e.g., images, videos, articles, etc., relating to the laws of thermodynamics. The speaker can, at their discretion, use material from the search results to display to the audience.

Search queries are formulated in a lookahead manner, e.g., the entire content of the speaker's material is scanned to determine appropriate search queries and search context. With speaker permission, oral statements made by the speaker can be used to derive search queries that are run in the background. With audience permission, questions asked by audience-members, e.g., students, can also be used to derive search queries in the context of the topic, and the results arising from such queries presented to both speaker and audience.

The transcription of oral statements to search queries is done by an artificial intelligence (AI) based speech-to-text module. Contextual formulation of search queries can also be done using AI or machine learning (ML) models. The AI or ML modules can reside on-device, e.g., on the computer used by the speaker, such that no data is transmitted to a server. Alternatively, the AI or ML modules can reside on a server, such that, with user permission, speech data (or transcript) and/or presentation material is sent to the server to determine search queries.

Search results are populated dynamically and grouped by search query, e.g., in the form of cards. For example, as shown in Fig. 1, results relating to the zeroth law of thermodynamics are grouped together; results relating to the first law grouped together; etc. The bottom bar is tabbed (106) by type of content, e.g., images, videos, audio, articles, etc. The speaker can pause the presentation at their discretion to use any of these: for example, articles can be used to show relevant studies to the audience; image results can be used to explain a relevant diagram; videos can be used to show animations; etc. Result-groups within the moving side (or bottom) bar are time-stamped (110a-c) to enable the speaker to easily navigate to particular topics within the side-bar based, e.g., on audience questions.

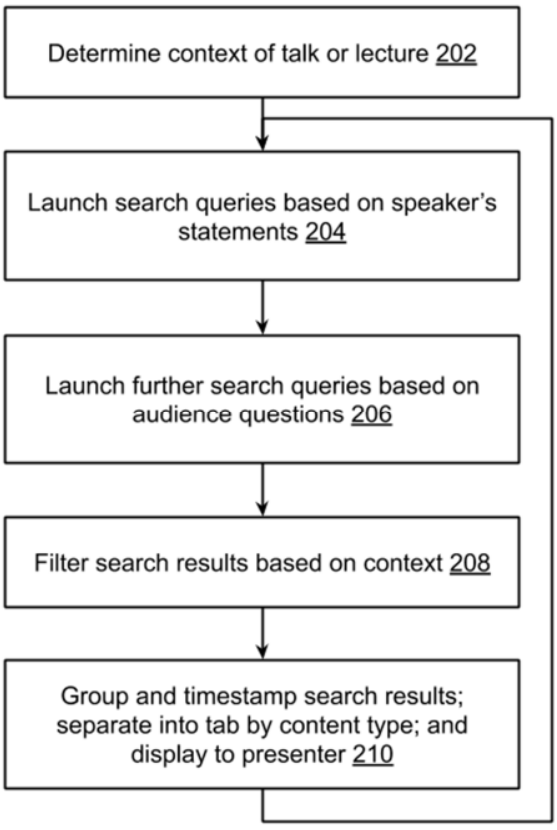

Fig. 2: Determining contextual search queries in an automatic, real-time, and dynamic manner

Fig. 2 illustrates an example procedure to determine contextual search queries in an automatic, real-time, and dynamic manner. With permission of the speaker, the context of the talk or lecture is determined (202), e.g., from the title or other contents of the presentation. Furthermore, details of the context can be determined by doing a lookahead scan of the content of the document or slide deck that is being presented.

With permission from the speaker, spoken statements made by the speaker, e.g., teacher or lecturer, are obtained and used to formulate queries to perform searches in the background (204). Furthermore, with prior permission, if an audience member, e.g., student, asks a question, that question can also be used as a query string in the context of the topic, and the search results can be presented both to the speaker and the audience (206).

The results of the searches are filtered based on the context (208). The results are populated dynamically in a sidebar or bottom-bar that is viewable by the speaker (210). For example, the results can be formed into a group of cards with respective timestamps, such that the speaker can navigate to the search results at a later time, e.g., to address questions raised by the audience. The results are tabbed based on content type, e.g., articles, images, videos, etc. At their discretion, the speaker can pause the presentation to use any of the search results. For example, articles can be used to show relevant studies to the audience; image results can be used to show a relevant diagram; videos can be used to show animations; etc.

In this manner, the techniques of this disclosure enable a speaker to be dynamically provided with relevant information and resources to deliver high-quality presentations, and to answer audience questions accurately and efficiently by referring in real time to relevant material obtained through contextual queries. The techniques enable a quick yet deep discussion of the topic on hand.

Further to the descriptions above, a user may be provided with controls allowing the user to make an election as to both if and when systems, programs, or features described herein may enable the collection or use of user information (e.g., information about a user's documents, spoken statements, social network, social actions or activities, profession, a user's preferences, or a user's current location), and if the user is sent content or communications from a server. In addition, certain data may be treated in one or more ways before it is stored or used, so that personally identifiable information is removed. For example, a user's identity may be treated so that no personally identifiable information can be determined for the user, or a user's geographic location may be generalized where location information is obtained (such as to a city, ZIP code, or state level), so that a particular location of a user cannot be determined. Thus, the user may have control over what information is collected about the user, how that information is used, and what information is provided to the user.

CONCLUSION

This disclosure describes techniques to assist speakers in a dynamic, real-time, and automatic manner. Speaker presentation materials and spoken statements are obtained with permission and are used to formulate contextual queries. Results of the queries that include contextually relevant content are grouped by topic and are displayed in a user interface, e.g., as a sidebar that is viewable by the speaker. The results can be displayed in a tabbed user interface that separates content based on content type.