

Technical Disclosure Commons

Defensive Publications Series

March 13, 2019

Providing readout of multimedia content in messages

Niyati Joshi

Vinh Ly

Okan Kolak

Follow this and additional works at: https://www.tdcommons.org/dpubs_series

Recommended Citation

Joshi, Niyati; Ly, Vinh; and Kolak, Okan, "Providing readout of multimedia content in messages", Technical Disclosure Commons, (March 13, 2019)

https://www.tdcommons.org/dpubs_series/2021



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

Providing readout of multimedia content in messages

ABSTRACT

This disclosure describes techniques to automatically extract information from messages that include multimedia content, e.g., image, video, etc., and provide a readout of the message content to the recipient. Extraction of the content can be performed using machine learning. For example, readouts can be provided for multimedia messages received via short message service (SMS), chat or messaging applications, email, etc. The techniques can be implemented as part of a virtual assistant application. The readout is provided upon user request. Messages are accessed and analyzed upon specific permission of the user.

KEYWORDS

- readout
- text-to-speech
- SMS
- chat message
- multimedia message
- virtual assistant
- object recognition

BACKGROUND

Users receive messages, e.g., SMS, chat, email, etc. on their mobile devices. Some mobile devices enable users to have the message read out to them, e.g., by providing a command to a virtual assistant application provided by the mobile device. For example, the user can provide the command upon receiving a message notification. However, currently, such readouts

are restricted to text messages (or text portion of multimedia messages). When a message includes an image, video, a file, a URL, etc., only the text portion is read out, or a simple readout such as “this message includes a photo” is provided. This makes for a less than satisfactory user experience.

DESCRIPTION

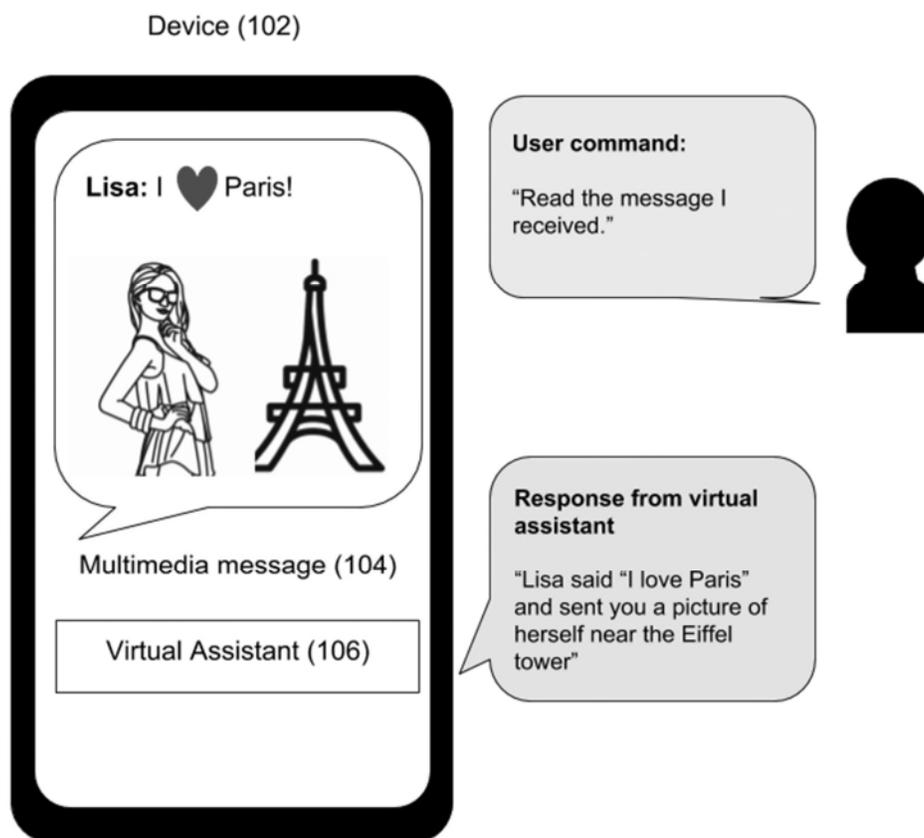


Fig. 1: Automatic extraction and readout of a multimedia message

Fig.1 illustrates an example of a virtual assistant application responding to a request from the user to read out a message that includes multimedia content. A user receives a multimedia message (104) on user device (102). In the example illustrated in Fig. 1, the message includes the text “I ♥ Paris” (where the ♥ may be text or emoji) and a photo of the sender (Lisa) near the Eiffel tower. With user permission, the message is analyzed, e.g., using on-device or server-

based mechanisms for object recognition and content/metadata extraction from media. The mechanisms can be accessed via an application programming interface (API).

When the user requests a readout of the message, e.g., to a virtual assistant (106) application on the user device, the extracted information is provided to the user. For example, the virtual assistant can respond “Lisa said “I love Paris” and sent you a picture of herself near the Eiffel tower” based on the object recognition techniques determining that the photo includes Lisa (if face recognition is permitted) and the Eiffel tower.

The described technique can be implemented as part of a virtual assistant application, a messaging application, an email application, device operating system etc., in devices that support messaging, e.g., smartphones, tablets, smartwatches, etc. The techniques are implemented with specific user permission. If the user denies permission to access their messages or restricts permission to specific messages, only such messages are accessed as permitted by the user. Users are provided with options to turn off automatic analysis and/or readouts of incoming messages.

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 collection of user information (e.g., information about a user’s 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 automatically extract information from messages that include multimedia content, e.g., image, video, etc., and provide a readout of the message content to the recipient. Extraction of the content can be performed using machine learning. For example, readouts can be provided for multimedia messages received via short message service (SMS), chat or messaging applications, email, etc. The techniques can be implemented as part of a virtual assistant application. The readout is provided upon user request. Messages are accessed and analyzed upon specific permission of the user.