Chat2Doc: From Chats to How-to Instructions, FAQ, and Reports

Britta Meixner
Centrum Wiskunde & Informatica &
FX Palo Alto Laboratory, Inc.
Amsterdam, NL
britta.meixner@cwi.nl

Matthew Lee
FX Palo Alto Laboratory, Inc.
Palo Alto, CA, USA
mattlee@fxpal.com

Scott Carter
FX Palo Alto Laboratory, Inc.
Palo Alto, CA, USA
carter@fxpal.com

ABSTRACT

Sharing multimedia via messaging apps is widely used. However, the timeline structure makes it difficult to retrieve content shared over time. It is not possible to organize accumulated knowledge so that it is concise for future use and easy access. So far, no system exists that combines the easy-to-use interface of a messaging app with a knowledge extraction system that can create multimedia documents and allows users to store and edit content for future use. In this paper, we propose a system that will enable individuals to collect, store, and automatically extract procedural knowledge from their messaging interactions. The system uses the well-known chat interface to communicate and adds the capability for users to tag text and media to organize content. It also adds a new thread-like structure to the previously only linear timeline of a chat. Knowledge from the chat can then be extracted into a high-quality multimedia document.

KEYWORDS

Hypermedia document; Chat; How-to instruction; FAQ; Document Generation; User Interface

1 INTRODUCTION

Chatting and messaging apps are not only the most commonly used apps on mobile devices, they also allow people to share a wide variety of multimedia using a simple, well-understood interaction metaphor of a conversational timeline. However, whereas existing messaging apps are well-suited for communicating and sharing content on-the-go, it is difficult to find and retrieve content generated and shared over time. Due to limited screen sizes and potentially many comments, information disappears at the top of the screen and is forgotten easily. Consequently, questions may remain unanswered or important information may escape one's memory because it is not possible to mark important elements or answers to questions directly in the chat.

Knowledge is generated while chatting over time, particularly the contextual and procedural knowledge on how to perform a task or how to solve a problem, that might be helpful for other people in the future. However, it is difficult to extract the most relevant content from a chat and organize it so that it is concise for future use and easy accessible by group members and (if desired) the general public. No systems exist so far that provide easy to use and well-known interfaces for the creation of multimedia documents which can be used as How-to instructions, answers to Frequently Asked Questions (FAQ), or printable reports from existing collections of group knowledge. Whereas knowledge management systems are able to deal with extensive amounts of knowledge, they might be too complex to create knowledge documents "on the side".

A system that will enable individuals to collect, store, and automatically extract procedural knowledge from their messaging interactions with collaborative groups is proposed in this paper. The system uses a chat interface as a basis to communicate and collect knowledge. Tagged text and media are used to organize content immediately and for future use. For example, messages can be marked as questions or to-dos. This adds new thread-like structures to the previously linear time-based structure of a chat. Knowledge from the chat can be extracted, edited, and enhanced. A web interface allows users to change the order of elements, change and add media, and create different playback paths for users with varied expertise. By leveraging how users naturally engage in group messaging, this new approach will result in a high-quality multimedia document that contains knowledge in an easily accessible way.

This work proposes a new approach for collaboratively and automatically generating navigable multimedia documents that organize procedural knowledge extracted from routine multimedia chat data from small task-oriented groups. Our proposed system can be a helpful tool for smaller task oriented groups. Exemplary groups that may benefit from our Chat2Doc system are people who jointly take care of an animal, groups that provide help to others or coordinate a common project, learning groups for exams, or members of a distributed team. These groups have the following in common: solving problems, especially when parts of the problem are repetitive tasks that may come up again in the future. There are certain issues that the group (repeatedly) has to deal with. The group may have discussions about topics leading to a good and practicable solution for the future. Usually, the communication in a group chat is about different topics and happens asynchronously because not all group members are available at the same time. This leads to a mix of topics, questions, and answers which makes it hard to find information at a later time. It can be helpful to extract the generated knowledge for future use as problems and tasks may reappear. Our system can help small groups from different areas.

In the paper, we first describe challenges in Section 2. After an analysis of related work in Section 3 our framework is described in Section 4. The paper ends with a conclusion and an outlook on future work in Section 5.

MultiEdTech'17, October 27, 2017, Mountain View, CA, USA

ACM ISBN 978-1-4503-5508-7/17/10…$15.00
https://doi.org/10.1145/3132390.3132392

2 CHALLENGES

While knowledge and content in the form of text, images, videos, and audios are created every day by millions of users in chats, they stay (rather) unconnected. Large collections of information are extended over time without filtering, structuring, or reusing their contents. However, no mechanisms exist, that allows easy organization and reuse of existing knowledge from a chat communication while solving a problem, and transform it into a better and more usable form like a How-to instruction, FAQ, or a printable report. For the creation of instructional and procedural multimedia documents, the use of multimedia is key, because many things can be better explained with a combination of text, image, animation, and video than just text. Considering the fact that these types of media are collected in chats every day while people communicate and collaborate in problem solving scenarios, it seems natural to mine these sources for appropriate content.

The system proposed in this work, Chat2Doc, does more than just search in a chat archive for information. In fact, it allows users to reuse and reorganize the chat contents for others. This requires new features in both, the chat app and the document creation system. Whereas the chat interface is excellent for collecting information, it is hard to find information in the original chat. Knowledge needs to be rearranged, edited, and extended in a document creation system. The system can also generate documents to target different user groups, such as beginners vs. experts or patients vs. doctors, creating even more complex hypermedia structures. A presentation-engine with a viewer is needed for displaying the resulting hypermedia documents. It should avoid information overload in a clear and easy-to-use interface. Designing such a system raises the following research questions:

- Extended chat:
  - Which features can help small task-oriented groups?
  - How can chats be structured to avoid a mix of topics?
- Document generation:
  - How can knowledge that was collected in group chat be transformed into a multimedia document?
  - How can users be supported in extending captured media?
  - How can media existing in the WWW be integrated?
  - How can a multimedia document be extended for different user groups/scenarios?
- Presentation engine with viewer:
  - What contents should be displayed in the resulting playback for a certain user?
  - How much user interaction is desired/necessary?

It is furthermore important to find out which types of knowledge and parts of conversations are users of a group willing to share with others.

3 RELATED WORK

We analyzed several commonly used messenger apps, business messenger apps, and work about the structuring of mails to identify useful features and "standard implementations" for Chat2Doc. We focus on chat interfaces that are widely used to collect information and knowledge.

Research in the area of email structuring can give hints on structuring chats/messaging apps. Mackay writes that "users archive messages for subject retrieval, prioritize messages to sequence work activities, and delegate tasks via mail" [4]. Millen et al. propose findings about a social bookmarking service and find out that there is a "willingness to share informational resources for the benefit of the organization" [7]. This work hints on the question of what knowledge users or groups are willing to share. Tang et al. conducted a limited field test on a mail system that provides message tagging, message threading, and email folders. An important finding is that "the interface has to communicate a clear model of threading, and especially help the user detect all new contributions to a thread" [12]. Whittaker et al. investigated email message re-finding. They found out that "people who received more threaded emails were more successful in their retrievals" [18]. Chat apps already provide a structure grouping messages by user or group. However, no hints could be found if a more fine grained structure in a conversation has the same effect as shown by Whittaker et al.

Messenger apps are available for mobile and desktop devices. Besides writing text messages, most of the apps support the upload of images, videos, and audios from the device. Emoticons or stickers/GIFs and group chats can also be considered as standard features. Most mobile apps support capturing (and then posting) images, videos, and audios as well. They also provide features to send the phones location or contact cards from the phones contact list. We analyzed Android apps with more than 10,000,000 users (on February 2, 2016) for features that could be helpful for group management (besides the already mentioned ones) and will only list the most useful features for our system. In addition to standard features, Skype [9] allows quoting of messages. WhatsApp [17] allows to mark messages which are then collected and displayed on a separate screen. WeChat [14] allows to add messages to favorites. Kakao Talk [3] allows marking of messages. Additional features are polls in groups, schedule for events in a group, and a photo and video album for group. Telegram [13] has a forward button for elements in the chat and a function to answer an element (which is cited and added at the bottom of the chat with the answer). Icq [2] also allows to quote messages which are then appended at the bottom of the chat with a reply. The more advanced features of Kakao Talk, Telegram and icq are the announcing of messages at the top of the screen and the quoting of messages which are then shown with the answer as new messages. Apps like WhatsApp, Skype, or Facebook Messenger [1] are available as both - mobile and desktop applications. None of the existing messaging apps have a threaded structure or allow people to mark questions and see directly linked answers. It is furthermore not possible to send time/location reminders.

Messaging apps for teams provide more advanced features than those for regular communications between friends. Besides group chats, files are stored with the chat and can be accessed anytime. All apps mentioned hereafter are available for desktop and mobile devices. All apps support a keyword search in a chat. Slack [10] allows to connect and synchronize several tools offering a variety of features. Besides searching in a chat, it is also possible to search in shared files which are indexed and archived once in the app. Messages can be marked for future reminders. Flock [8] allows the creation of polls and to-dos from the desktop app. Threads to follow up on topics can be created in Flock and Slack and are presented in

a separate section outside the originating chat. Crew [11] allows to send tasks to group members and notifies when a task is finished.

Team and business messaging apps described in this section provide a large number of features that are useful in the daily worklife. However, considering our target users of small, mostly private groups, it has to be stated that regular chats as described in the previous section provide no time or task management features. The apps described in this sections in contrast provide all features but may lead to confusion for inexperienced users trying to take care of simple managements task with the given number of features. While it is possible to search for contents (even in shared files) and pin them, it is not possible to extract knowledge for future use.

Multimedia documents can be created in different ways. In the following we focus on mobile and collaborative document creation. An overview of traditional authoring tools for hypermedia documents like hypervideos or multimedia presentations can be found in Meixner [5]. Weir proposes a "workflow that encourages users to contribute and refine subgoals for a given how-to video" [16]. Then Weir et al. deployed a "live website with a workflow for constructing subgoal labels" [15]. The system is tested on a set of web programming videos. Video segments are labeled by users "while answering reflective summary questions", then other users can vote on labels that best describe a subgoal from the first step. In a third step, the result is proofread and can be corrected. The tool allows the collaborative creation of an outline for non-linear navigation in How-to videos. A modification of this process can be used in our system to collaboratively select content from the chats. It can also help in the viewer to improve documents.

4 FRAMEWORK

The Chat2Doc system consists of four major components (see Figure 1, left): the extended multimedia chat app (A), the multimedia document generator (B), and the multimedia document presentation engine with viewers (C) and report generation (D). The extended multimedia chat app has new features that are helpful for small task-oriented groups and will be available for mobile devices and desktop computers. The document generator allows the creation of multimedia documents like How-to instructions, FAQs, or reports. The multimedia document presentation engine with viewers is used to filter, view, and play the documents. They may be exported for print.

4.1 Extended Multimedia Chat Application

Most of the regular chat apps have the following features: sending text, images, videos, and audio messages, edit images and videos, send animated gifs and emoticons, send contact cards, and communicate with groups. Our app (described in more detail in [6]) has additional features (listed below) designed for small task-oriented groups. These elements are created in the chat view during a conversation. Depending on the chat entry, a dialog may open, where the user can input additional information (such as due-dates, selected group members, or durations for validity). The created element is then marked in the chat and/or shown on the extended start/overview screen of the app, which not only shows the groups and past chats, but also unanswered questions and other elements

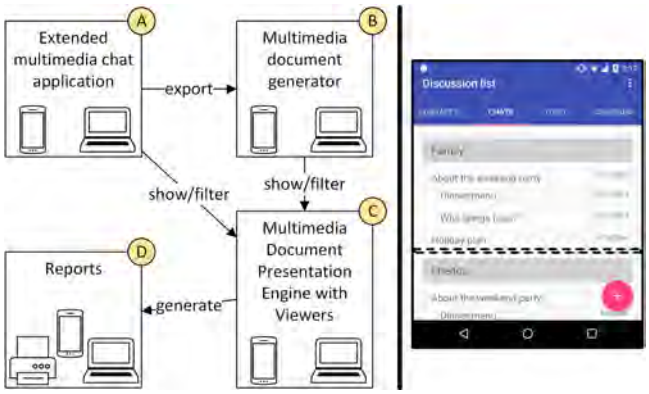

Figure 1: Overview of the system (left) and start screen of the chat app (right)

that may require an action. A preliminary version of the start screen from our app is shown in Figure 1 (right):

- **Questions:** Questions are highlighted in the chat and appear on the start page until they are marked as answered (e.g. Figure 1 (right), "Who brings food?"). They start a thread-like structure of elements breaking up the timeline-based view of a chat ([removed for blind review]).
- **Notes:** Notes are highlighted in the chat and have to be marked by the recipient to hide it from his/her start page.
- **To-Dos/Tasks:** Tasks are highlighted in the chat and appear in a to-do list in the app. They have to be marked as done by at least one of the assigned group members, may have a deadline, and may require feedback (like a photo of the result).
- **Message to future:** Messages can be scheduled to be sent at a future time and be delivered only then.
- **Time-based reminder:** This reminder is created like a regular message but is assigned a reminder date/time. The reminder immediately appears in the timeline list of messages, but reappears as the most recent message at the set date/time.
- **Location-based reminder:** Similar to the time-based reminder, but the sender sets a reminder location instead of a date/time for the message to reappear.
- **Time- and location-based reminder:** Similar to the time-based reminder, but the sender also sets a reminder location for the message to reappear.
- **Group calendar:** A group calendar gives an overview of personal and group activities.
- **Mark messages:** Messages that were posted to the chat timeline can be marked as important or "liked".
- **Search:** Search for a particular word or for images/ audio/video in a chat. Media search requires advanced search methods like search by date, search by media type, search by content, etc.

Each of the features described above can be sent to the whole group, a part of the group, or a single group member.

4.2 Document Generator

The newly introduced features from the previous section add an additional annotation level to the previously linear chat. Different

ways of marking elements hint on their importance. However, it can be helpful to not only mark elements but also extract a semantically connected set of elements containing knowledge from the chat and transform it into a more readable form. This can be done as follows:

User-guided: In the user-guided creation, an author has to perform several steps. First, chat elements have to be selected. This process starts with the user clicking one chat entry, then a selection tool provides a list of the following elements in the chat. After that, the user can choose for each element, if it should be added to the document or not. Then, selected elements are copied to the editor area where they can be rearranged, edited, and merged with the previous or following elements. New elements can be added from the local file system or the Web if desired. Formatting can be applied. In the end, the document is exported.

Semi-automated: In the semi-automated creation, several algorithms are applied after the user selects from where in the chat to start the selection. Basic text analysis as well as the analysis of annotations (likes, markers, etc.) suggest relevant entries, which may then be altered by the user. During the copy process, unnecessary parts of the texts (like emoticons or unnecessary spaces) are removed and elements users posted as consecutive smaller messages are merged to one text. The resulting list of elements can then be edited as described for the user-guided process. Here, the user is supported by retrieval and search algorithms that provide video scenes, images, or other materials from the Web.

Automated: The automated creation learns from both other modes and automatically exports the desired document format using learning algorithms.

4.3 Multimedia Document Presentation Engine

Our multimedia document viewer provides features to display all types of media as well as elements for navigation. The viewer adapts to the playback device and the characteristics of the user. Depending on the playback device, the amount of data can be adjusted as well as the sizes of the media elements. Slower connections may focus on text and images, while videos are shown using higher bandwidths (in case different types of media are available). Navigational elements allow users to skip parts or select parts with more or less information. Further filtering and adaptation can be applied based on the user characteristics which may be provided at the beginning of the playback. Furthermore, functions are available that allow users to extend the currently viewed multimedia document if information seems to be missing or presented in a wrong way, enabling previously passive viewers to become collaborating authors. Mechanisms include peer reviewing and voting of content to ensure a good quality of newly added information.

5 CONCLUSIONS

In this paper we described the Chat2Doc system, which combines the easy-to-use interface of a messaging app with a knowledge extraction system that can create multimedia documents. This allows users to store and edit content for future use as How-To Instructions, FAQs, or summary reports. The Chat2Doc system will enable individuals to collect, store, and automatically extract procedural knowledge from their messaging interactions with others. The extended chat interface allows tagging of text and media to organize content, a new thread-like structure extending the linear timeline, and provides usable features like to-do lists or calendars. Generated knowledge can be extracted automatically from the chat into a high-quality multimedia document. Accordingly, created hypermedia documents not only compose information in a meaningful way, but also provide adaptation capabilities using the underlying hyperlink structure. Our review of literature and apps showed that no system exists that combines the aforementioned features.

This paper describes work in progress. User tests and the usage of the system have to show which features are used, how, and for what. A critical factor of the system is collecting enough knowledge that allows manual and semi-automated creation of documents. This is necessary to train the system for the automated mode.

ACKNOWLEDGMENTS

Parts of this work were carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship Programme.

REFERENCES

[1] Facebook. 2016. Messenger (57.0.0.31.81). Website https://www.messenger.com/ (Feb. 17, 2016). (2016).
[2] ICQ LLC. 2016. icq (6.5). Website https://icq.com/android/en (Feb. 17, 2016). (2016).
[3] Kakao Corp. 2016. kakaotalk - Welcom to the world of Kakao Talk (5.3.3). Website http://www.kakao.com/talk (Feb. 17, 2016). (2016).
[4] Wendy E. Mackay. 1988. Diversity in the Use of Electronic Mail: A Preliminary Inquiry. *ACM Trans. Inf. Syst.* 6, 4 (Oct. 1988), 380–397.
[5] Britta Meixner. 2014. *Annotated Interactive Non-linear Video - Software Suite, Download and Cache Management.* PhD Thesis. Universität Passau. https://opus4.kobv.de/opus4-uni-passau/files/222/Meixner_Britta.pdf
[6] Britta Meixner, Matthew Lee, and Scott Carter. 2017. Managing Family Healthcare with Multimedia Chat Apps: A Survey on What is Missing. In *MMHealth'17: The 2nd International Workshop on Multimedia for Personal Health and Health Care.* ACM, New York, NY, USA.
[7] David R. Millen, Jonathan Feinberg, and Bernard Kerr. 2006. Dogear: Social Bookmarking in the Enterprise. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI '06).* ACM, New York, NY, USA, 111–120.
[8] Riva FZC. 2016. Flock - Free group & 1-on-1 chat for your team. Website https://www.flock.co/ (Feb. 17, 2016). (2016).
[9] Skype and/or Microsoft. 2016. Skype - Skype keeps the world talking, for free. (6.20.0.618). Website https://www.skype.com/en/ (Feb. 17, 2016). (2016).
[10] Slack Technologies. 2016. slack - A messaging app for teams who put robots on Mars!! Website https://slack.com/ (May 17, 2016). (2016).
[11] Speramus, Inc. 2016. Crew - Communicate with teams and managers on the go. Website https://crewapp.com/ (May 17, 2016). (2016).
[12] John C. Tang, Eric Wilcox, Julian A. Cerruti, Hernan Badenes, Stefan Nusser, and Jerald Schoudt. 2008. Tag-it, Snag-it, or Bag-it: Combining Tags, Threads, and Folders in e-Mail. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08).* ACM, New York, NY, USA, 2179–2194.
[13] Telegram. 2016. Telegram - a new area of messaging (3.5.1). Website https://telegram.org/ (Feb. 17, 2016). (2016).
[14] Tencent Inc. 2016. WeChat (6.3.13.64_r4488992). Website https://web.wechat.com/ (Feb. 17, 2016). (2016).
[15] Sarah Weir, Juho Kim, Krzysztof Z. Gajos, and Robert C. Miller. 2015. Learnersourcing Subgoal Labels for How-to Videos. In *Proc. of the 18th ACM Conf. on Computer Supported Cooperative Work & Social Computing (CSCW '15).* ACM, New York, NY, USA, 405–416.
[16] Sarah A. Weir. 2014. Learnersourcing Subgoal Labels for How-to Videos. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems (CHI EA '14).* ACM, New York, NY, USA, 945–950.
[17] WhatsApp Inc. 2016. WhatsApp - Simple. Personal. Real Time Messaging. (2.12.449). Website https://www.whatsapp.com/ (Feb. 17, 2016). (2016).
[18] Steve Whittaker, Tara Matthews, Julian Cerruti, Hernan Badenes, and John Tang. 2011. Am I Wasting My Time Organizing Email?: A Study of Email Refinding. In *Proc. of the SIGCHI Conf. on Human Factors in Computing Systems (CHI '11).* New York, NY, USA, 3449–3458.