
LifeSampler: Enabling Conversational Video Documentary

Ryan P Spicer

Research Assistant
Reflective Living
Arts Media and Engineering
Arizona State University
699 South Mill Ave #393
Tempe, AZ 85281 USA
Ryan.Spicer@asu.edu

Aisling Kelliher

Assistant Professor
Reflective Living
Arts Media and Engineering
Arizona State University
699 South Mill Ave #396
Tempe, AZ 85281 USA
Aisling.Kelliher@asu.edu

Copyright is held by the author/owner(s).
CHI 2008, April 5 - April 10, 2008, Florence, Italy
ACM 978-1-60558-012-08/08/04.

Abstract

Today's multidisciplinary, fast-paced and innovative workplaces present new challenges in facilitating effective communication between diverse team members and ensuring successful transfer of knowledge within a flexible workforce. In this paper, we present Conversational Documentary, a model for supporting constructive audiovisual dialog between workplace colleagues that also aims to archive and interpret the work practices and approaches of a creative community. We discuss the development and initial evaluation of the LifeSampler, a prototype audiovisual system designed to support and test our model, and propose directions for future development based on our preliminary results.

Keywords

User-generated Video, Conversational Documentary, Lifestreaming

ACM Classification Keywords

H.5.1 [Multimedia Information Systems]: Video, H.5.3 [Group and Organization Interfaces]: Computer-supported cooperative work, H.5.2 [User Interfaces]: Screen Design, H.5.3 [Group and Organization Interfaces]: Asynchronous Interaction.

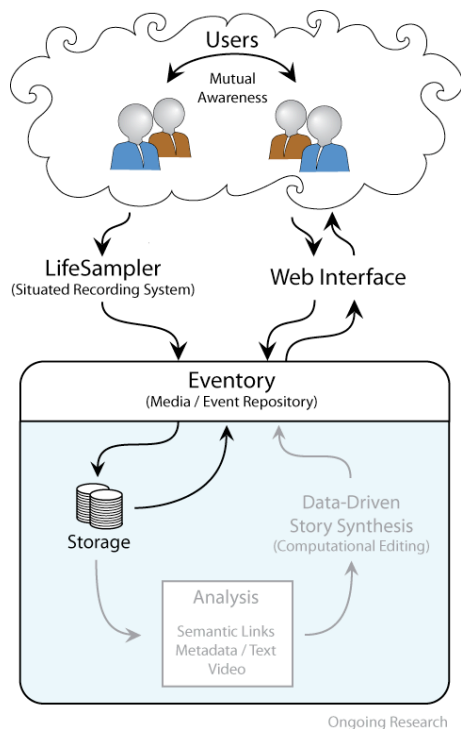


Figure 1 - Conversational Documentary Ecosystem

Introduction

Transdisciplinary teams face challenges both in achieving effective communication and mutual awareness among diverse groups of practitioners and in promoting knowledge transfer among team members as the group's composition changes over time. We propose a model of collaborative video-making, *Conversational Documentary*, which emphasizes both the asynchronous conversations generated through users' contributions of media and the long-term archival and knowledge-transfer potential of the resulting video [Figure 1]. Additionally, we evaluate the *LifeSampler*, a prototype situated system implementing the Conversational Documentary Model.

The LifeSampler is a physically situated video recorder designed to capture short interview-style segments of creative practitioners discussing their work in response to questions submitted by other users. The system uses a very simple user interface and automates many repetitive, time-consuming processes to lower the considerable barrier to entry faced by newcomers to the video production world.

Video and usage information from the LifeSampler are uploaded automatically to the *Eventory*, an online media repository where users can retrieve content, annotate media with typed semantic links, and explore user activity visualizations [1]. Synthesis of the resulting media into narratives based on computational analysis and user input creates longer-form content suitable for archival and knowledge-transfer purposes.

Summary of Prior Work

The popularity of social networking sites like Twitter [www.twitter.com] and Facebook's "status" widget

[www.facebook.com] suggest that certain communities of users are already motivated to publish constrained text-based "status updates" to improve mutual awareness in their social circles. This usage, however, occurs in a casual, social context.

Automation of video production processes from recording and capture [2] to editing [3,4] has been the topic of numerous research efforts. These automated systems rely on metadata input by skilled users during the recording [5] and/or logging [6] processes. Some systems generate metadata based on sensor data captured at record-time [7] or computational analysis during logging [8,9]. However, these systems are constrained by reliance on controlled conditions, the use of external hardware, or operation by skilled users.

In the commercial arena, a variety of tools offer video recording and distribution. These services range from broadcasting tools like YouTube's Video Uploader and "lifestreaming" service Justin.tv, to blog-like tools such as Facebook's video-wall application and Vimeo [www.vimeo.com].

With the exception of the Facebook tool, these systems focus on the one-to-many case, where a small group of creators generate media that will be received by a large, mostly passive audience. Viewers do not have the opportunity to respond in kind to an equally-sized audience. The Conversational Documentary model closes the feedback loop more directly, creating many-to-many relationships where viewers are also producers and media-production becomes a conversation.

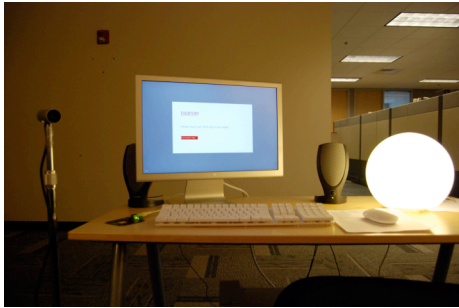


Figure 2 - The Prototype LifeSampler Hardware

Motivation

We propose that the Conversational Documentary model may be best realized in a system that builds upon the most useful characteristics of each of these existing classes of system. Twitter and similar services afford short, spontaneous text updates that encourage asynchronous sharing of information. The expressive qualities of video allow the transfer of large amounts of information in a short time. We add specific, intentional probes to encourage deeper discourse and prompt users to contribute to specific goals.

The LifeSampler system builds on Twitter's single "what are you doing?" question with a variety of more focused probes and allows users to pose their own questions to the community. These questions are intended to provoke deeper reflection upon creative activity than a simple "what are you doing?" and to allow the user community to shape the discourse. The system is initially seeded with questions relevant to the area of work, in the case of the preliminary study, "what does your research tell us about you?" and "how do you believe we can improve mutual awareness non-computationally?". Users may also contribute their own probes, which are then presented to other users along with the preloaded questions.

Video provides an information-rich, approachable method for creators to demonstrate their work and solicit feedback. It requires less compositional energy than text and provides more information and engagement than audio alone. High-quality video production, however, presents a very high barrier to entry, with complex tools and techniques to be mastered at all stages of production. By designing a situated system in which the technically demanding

aspects of setup are performed once, by an expert user, and automated software handles the time-consuming work of capturing and compressing video, we hope to lower this barrier to entry and encourage users from all background to contribute their expertise and opinions in the form of short video clips. These practitioners become storytellers sharing their unique perspective and process with members of a broader community who can then respond in kind.

Within the Conversational Documentary model, individuals are encouraged to produce short video clips or segments in response to each other's prior work. This literal video conversation encourages mutual reflection and awareness not only among those creating the video content, but also among all those viewing the video conversation. In addition, we intend to develop computational editing tools that assemble these video conversations into data-driven story sequences for review by future members of the community. In this way, we hope that Conversational Documentary facilitates both short and long-term communications and knowledge transfer within organizations and communities.

Prototype System Details

The LifeSampler platform consists of a Macintosh G5 workstation, a firewire video camera, an RFID reader, and the Java-based LifeSampler software [Figure 2]. The kiosk software communicates with an event-focused media repository server, the Eventory [9]. The hardware is positioned in an informal space, which affords users a casual but public interaction with the system. We hope that situating the recording environment in public clearly defines the expected use of recorded content and avoids creating an

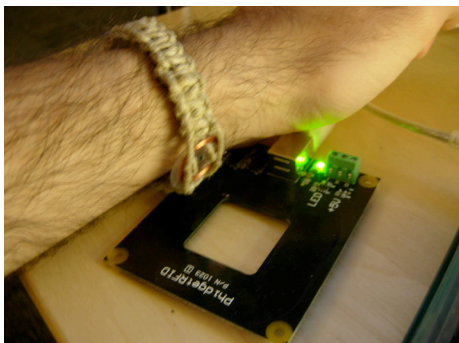


Figure 3 - A customized RFID Tag Bracelet and the RFID Reader

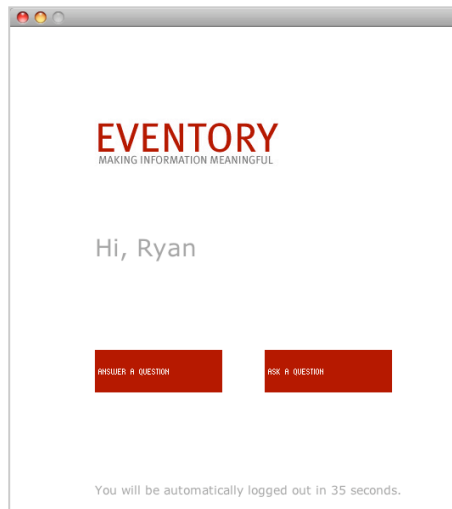


Figure 4 - LifeSampler UI, Choose to ask or answer a question

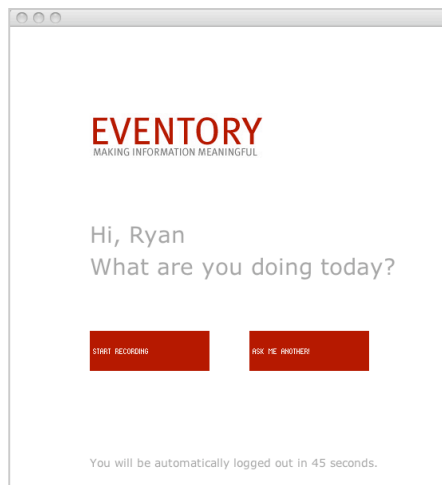


Figure 5 - LifeSampler UI, Record or Ask Me Another

unwarranted expectation of privacy for users, as one might feel in, for instance, a closed office.

The system is designed for extreme simplicity – at no time does the user have more than two choices, each of which is represented by a large text-based button. Our intent is to lower the barrier to entry – capture, compression and metadata capture are all handled by the system, with sensible defaults pre-configured.

In a typical interaction, the user identifies herself using a wearable, customizable RFID bracelet [Figure 3]. The user is then presented with two large text buttons – “Answer a Question?” or “Ask a Question?” [Figure 4]. In the former case a question in the form of a text probe or a brief video clip is displayed; if the user is not interested in answering this question, she may opt to skip it and another probe will be displayed [Figure 5]. In the latter case, she is prompted to prepare her question. She then uses onscreen controls to start and stop the video recording. The LifeSampler limits each response to one minute to encourage users to focus their thoughts and to keep answers in short chunks suitable for computational editing. Once recording is complete, the user reviews the clip and may choose to discard it or to submit it to the Eventory. The system compresses the video to web-friendly size and format and uploads the resulting file. The user’s name, the date and time, the location of the kiosk, and the question are automatically added to the file’s metadata.

Once an answer has been uploaded to the Eventory, both the original user and others may view the video. Users may also create typed semantic links between the video and other entities in the Eventory including users, media files and events. For example, a video

might “include” Jim and “be inspired by” a group meeting. This integrates the video file into a web of meaning that may be leveraged both by users and computational tools to infer meaning in the video content. Typed links and their relevance to meaning creation are discussed in more detail in [1].

Preliminary Results

We have conducted a preliminary user-study to gauge the effectiveness of the prototype system and generate user-centered goals for future development. Five users, four from within the research group and one from another research group in our department, were asked to make daily use of the LifeSampler and Eventory system for five days in support of their research work and to report on their experience. The system recorded usage statistics on all aspects of their use, including the questions they opted to skip and to answer, when they logged in, when they timed out, and so on.

Two study participants became power-users of the system, submitting the majority of content, while the others remained occasional contributors. Over the course of the study, these users contributed 22 new video questions and 38 responses to questions. This behavior suggests that users are at least as interested in soliciting feedback as they are at giving it.

One of the preloaded questions with which the system was seeded asked users to suggest improvements which prompted the following requests:

- Provide a mechanism for users to bookmark and return to challenging or interesting questions after thinking about them.

- Allow users to run the software from their own workstations rather than traveling to a physical installation.

In a survey administered at the end of the study period, several of the users expressed a desire to choose the question to be answered from a list of questions, rather than having it randomly selected. One user expressed frustration at having heard another user discuss a question he had contributed, which she wanted to answer but was unable to locate. Logs indicate that several users executed similar patterns of behavior – choosing “ask me another” several times in sequence while hoping for a specific question to be displayed. This behavior indicates that goal-oriented users might benefit from a search/browsing interface to rapidly select a specific question.

Though the preliminary study was brief, feedback from all participants indicates that participation in the study enhanced their awareness of and familiarity with their co-participants. The questions and answers served as an “ice-breaker” of sorts, promoting further discussion outside the Conversational Documentary framework.

Future Work

Situated LifeSampler

In order to maintain the simple user interface of the existing prototype, we do not wish to discard the general concept that a user will be presented with a system-selected initial probe. In the future, however, we intend to analyze usage patterns to select the questions that are presented to the user. For instance, newly-contributed questions and those apparently of interest to other community members (as measured by

the frequency of submitted answers) may appear more frequently than those often skipped.

We plan to implement a browsing/search system by which users may locate specific questions they wish to answer. This mode of interaction will be available from a button on the question-prompt screen. In this way, we preserve the simplicity of the existing interface while providing goal-oriented users with options.

Web-Based LifeSampler System

In response to users’ desire to contribute video from their own systems, one of our main priorities is the development of a web-based LifeSampler, integrated directly into the Eventory user interface. While this fundamentally alters the user experience, particularly the public context for recording, it should allow more participation by users who are unable to visit the physical installation.

In both the situated and web-based LifeSampler systems, we hope to provide a mechanism for users to respond directly to other users’ video contributions. Though this functionality exists through the Eventory’s semantic link functionality, we wish to integrate it with the recording system so that a user can explicitly indicate from within a single UI, for instance, “this response is inspired by another user’s response to a similar question.”

Data-Driven Story Synthesis and Authoring

Building on our prior work in online rich-media storytelling [10], we plan to develop computational methods for synthesizing meaningful stories from the contents of the contributed videos. Simple examples of this type of story might include a sequence showing the

evolution of one user or a community of users' responses to a single question over time, or a sequence presenting the interplay between questions and answers provided by users in response to one another. We hope that the computational system may, in this way, become a partner with the users in capturing content to generate relevant narratives.

The data to power this story-synthesis will be drawn from both the metadata captured and annotated by the LifeSampler system, and by the typed semantic links that users may create among content within the Eventory system [1]. The structure of these links may be analyzed to infer meaningful relationships among content items that may then drive the synthesis of story structures.

Conclusions

The Conversational Documentary model and LifeSampler system provide the tools needed to establish and strengthen communication among transdisciplinary teams. The LifeSampler lowers the barrier to entry in creating and sharing video content by providing a simple interface that minimizes complex decisions on the part of users. The system also archives data and provides metadata for computational editing tasks, the products of which will support knowledge transfer within the user community by allowing new members to browse the compiled experience of current and previous community-members.

Acknowledgements

We thank the students, faculty and staff of Arts, Media and Engineering, particularly those in the Reflective

Living subgroup, for their assistance in the development of this system.

Citations

- [1] Wang, X., Mamadgi, S., et al. Eventory--An Event Based Media Repository. ICSC 2007 Proceedings, 17-19 September, 2007, Irvine, California (2007) 95-104
- [2] Barry, B. Mindful Documentary. Massachusetts: Massachusetts Institute of Technology, 2005.
- [3] Bocconi, S. Vox Populi: Generating video documentaries from semantically annotated media repositories. Eindhoven: Technische Universiteit Eindhoven, 2006.
- [4] Casares, J., Long, A.C., et al. Simplifying video editing using Metadata. Proceedings of DIS, 2002.
- [5] Nack, F, and Putz, W. Designing Annotation Before It's Needed. ACM Multimedia. 2001.
- [6] Diakopoulous, N. and Essa, I. Videotater: An Approach for Pen-Based Digital Video Segmentation and Tagging. Proceedings of the 19th annual ACM symposium on User interface software and technology. 2006.
- [7] Liu X., et al. SEVA: Sensor Enhanced Video Annotation. Proceedings of ACM MM, 2006.
- [8] Song, Y., et al. Semi-Autonomous Video Semantic Annotation Based on Active Learning. Proceedings of SPIE, 2005.
- [9] Vezanni, R., et al. A semi-autonomous video annotation tool with MPEG-7 content collections. Proceedings of IEEE ISM, 2006
- [10] Kelliher, A. & Davenport, G. Everyday storytelling: supporting the mediated expression of online personal testimony. Proceedings HCI, 2007.