

Interaction Design for the Exchange of Media Organized in Terms of Complex Events

Anthony Jameson and Sven Buschbeck*

DFKI, German Research Institute for Artificial Intelligence
Saarbrücken, Germany

Abstract. Even the most sophisticated automatic recognition of events must often be paired with an appropriate design of the users' interaction with those events. This paper presents three presumably typical use cases and associated interaction design proposals, which illustrate (a) how untrained users can benefit from the organization of media in terms of complex events; (b) how they can have their own media categorized in this way without having to invest much effort; and (c) how they can even create complex event instances with novel structures, without having to think explicitly about event structures.

1 Introduction

As will be shown by many of the papers that will be presented at the EVENTS 2010 workshop, the automatic identification and processing of events raises many technical challenges. But even before solutions to these problems have been found, we have to consider exactly how people might interact with systems that make use of representations of events. Having a clear idea of use cases, scenarios, and interaction designs can help us to see which technical problems are most important and what requirements need to be met.

This workshop paper considers how the recognition and representation of events can enhance interaction in a particular type of system: a media marketplace in which professional and amateur users contribute and exchange various types of media, most typically photos and videos (but also other types, such as audio files and text documents). One underlying idea is that it is often helpful for such media to be indexed and organized in terms of events that they depict or describe, in addition to more familiar indexing on the basis of time, location, tags, and named entities (such as people).

More specifically, we consider how interaction in such a marketplace can be enhanced if not only atomic events but also *complex events* are represented: Such an event may extend over a considerable period of time and consist of subevents, some of which in turn may be complex events. A simple example of a complex event is a soccer tournament, which comprises two or more rounds and a number of games, each of which can in turn be viewed as a complex event.

* The research described in this position paper is being conducted in the context of the 7th Framework EU Integrating Project *GLOCAL: Event-based Retrieval of Networked Media* (<http://www.glocal-project.eu/>) under grant agreement 248984.

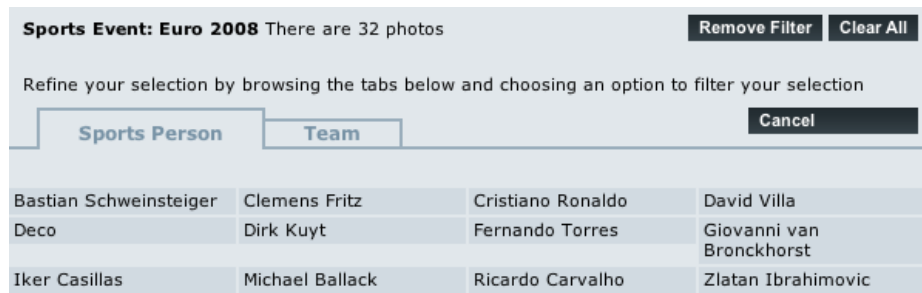


Fig. 1. Partial screenshot from the photo exchange site Sport Photo Gallery (<http://www.sportphotogallery.com/>). Though the site offers many photos of the Euro 2008 tournament, it is not possible to navigate among them in terms of the structure of the tournament.

We will present several scenarios and interaction designs that should help to stimulate thought on the following questions:

1. How could users benefit from the representation in the system of complex events, as opposed to having only simple events represented?
2. How can a user and a system collaborate to build up and maintain a representation of complex events, without any requirement for users to invest more than a minimal amount of effort?

This work is being done in the context of the integrating project GLOCAL.¹

2 Why Do We Need Complex Events?

Suppose you are an (amateur or professional) photographer or journalist who wants to share, buy, or sell media about the first half of the final game of the 2008 European Cup soccer tournament. Media concerning this event can be found in a number of media exchange sites, including Flickr.²

Citizenside.com³ is an example of a site that specifically supports selling of the media by amateur photographers to professional organizations, such as news agencies. Although this site organizes and indexes media in quite sophisticated ways, you would run into difficulty if you wanted to think in terms of parts of particular tournaments: The site does not organize media in terms of complex events like tournaments.

In the Sport Photo Gallery site,⁴ which is dedicated to sports photos (Figure 1), you can find the “Event” Euro 2008, but the media about it are indexed only in terms of players and teams, not parts of the tournament.

¹ Since a special session of the EVENTS 2010 workshop is being devoted to this project, we assume that the workshop proceedings will contain an introductory overview of the project; therefore, we do not include such an overview in this submission. If necessary, we can add such an overview in the final version of this paper.

² <http://www.flickr.com/>

³ <http://www.citizenside.com/en/sell-share-photos-videos.html>

⁴ <http://www.sportphotogallery.com/>

It may help to look at this absence of complex events in terms of an analogy: The way in which photos and videos can be embedded in a Google Map—say, of Athens—shows that it is feasible and useful to organize media in terms of a large, coherent structure—in this case, the map of a city. But suppose that some of these media concern events at a conference—for example, a talk in a session of EVENTS 2010, which is in turn a subevent of SETN 2010. Google Maps can show the conference building, but it has no way of representing the additional dimension: the structure of the “conference event”.

3 Use Case A: Navigating Via Event Structures

Suppose now that we have a media marketplace that includes:

- structures for complex events;
- media attached to particular events.

(We will discuss in below how the structures and the media will get into the system.)

Then a user can:

- 1. . . . find a complex event with some combination of keyword search, use of a map and a calendar, and/or providing an example medium about that event; Although finding an optimal interaction design for this sort of event search is an interesting challenge, it is not very difficult to find an acceptable solution, so we do not provide any concrete examples in this paper.
- 2. . . . navigate down the hierarchical structure of the complex event to find the part that they are interested in. One way of allowing this sort of navigation is to visualize the complex event as a tree structure in which each node represents an event or a subevent. In the hypothetical screen Figure 2, the user is focusing on the node for the subevent “first half of the final game”, and the media associated with that subevent are shown on the right-hand side of the screen. Nodes representing higher-level events can also have media associated with them, for example a video that covers the entire game.⁵

4 Use Case B: Inserting New Media Into an Event Structure

Even if we grant that users could benefit from this type of organization, the question arises of how media are going to get organized in this way. Realistically speaking, we cannot expect most users to spend a lot of time carefully creating complex event structures and assigning media to particular parts of these structures. So on the one hand, we need system-side processing that can handle a lot of the work of creating and populating complex event structures. On the other hand, since we cannot assume that a

⁵ The visualizations this paper were created with the MindManager software; they therefore do not reflect the appearance of the interfaces that will ultimately appear in the GLOCAL system.

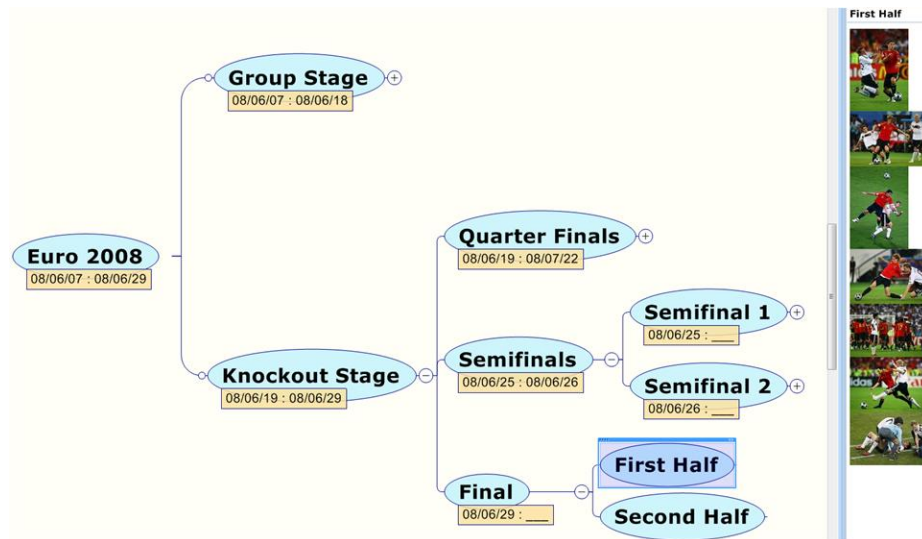


Fig. 2. Proposed visualization of the structure of a complex event in such a way that it can be used for browsing for media associated with subevents.

fully automatic solution will be satisfactory, we have to design the user interaction in such a way that users can help the system out without investing much effort.

In this use case, we consider how users might insert media into an existing complex event structure. (The problem of creating such a structure in the first place will be considered below.)

Suppose, concretely, that a photographer has created photos and videos of the Euro 2008 final and would like to add them to the Glocal site (e.g., to sell them or to share them with friends).

In Figure 3, she opens up a new node “New Media” under the “Final” event and uploads the media to the space on the right (which serves as a sort of inbox).

The user could in principle specify by hand whether each medium belongs to the first half, the second half, or the whole game (as with a video that includes highlights from both halves). But the system should be able to do this work largely automatically. Essentially, it can compare the space and time coordinates of the new media—and the low-level properties of their images—with those of the already categorized media.

In Figure 4, the left-hand side of the screenshot shows the system’s tentative sorting of the images. The small blue and white icons indicate the system’s confidence level: the more blue, the higher the confidence.

The right-hand side of the screenshot shows why it can be important to leave the last word to the user: The user has now deleted two of the low-confidence images (which she now recognizes as being largely irrelevant) and accepted the system’s classification of the other images. This example illustrates that, if the user can count on a reasonable amount of intelligence on the part of the system, the user can save some of her own time, even if the system’s performance is imperfect. With a bit of effort, the user could have recognized by herself that the photos of the team lining up before the game and

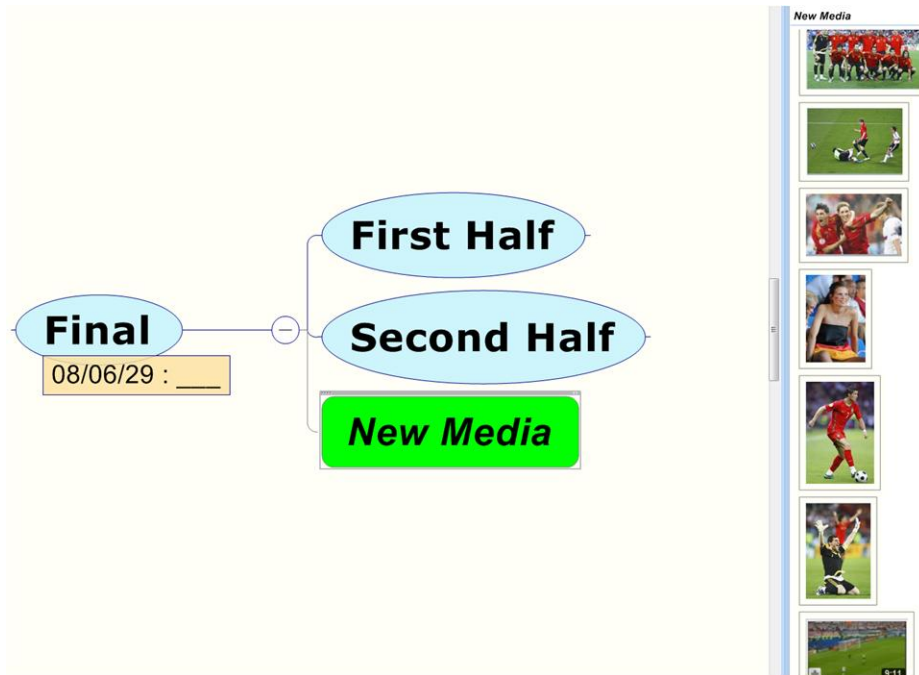


Fig. 3. Hypothetical screenshot of a situation in which a user is preparing to insert a number of new media into the structure for a complex event.

of the young lady in the stands do not really belong in the same category as the other photos and videos. But if she knows that the system will make it easy for her to remove any superfluous photos, she doesn't have to be so selective when offering them in the first place.

5 Use Case C: Creating a New Complex Event

But what if the user's new media concern a complex event that is not already represented in the system—maybe because it is of only local interest?

Specifically, assume that a mother has taken photos and videos of her 14-year-old daughter's local soccer tournament. The user will have to create a new complex event instance with an appropriate structure. So in principle, she needs either to find an existing event structure that she can instantiate or create a (partially) new structure that is suitable for describing her event.

The main challenge lies in the fact that most users won't be willing or able to reason in terms of event structures.

The approach that we propose is to support a "copy, paste, and modify" style of event creation.

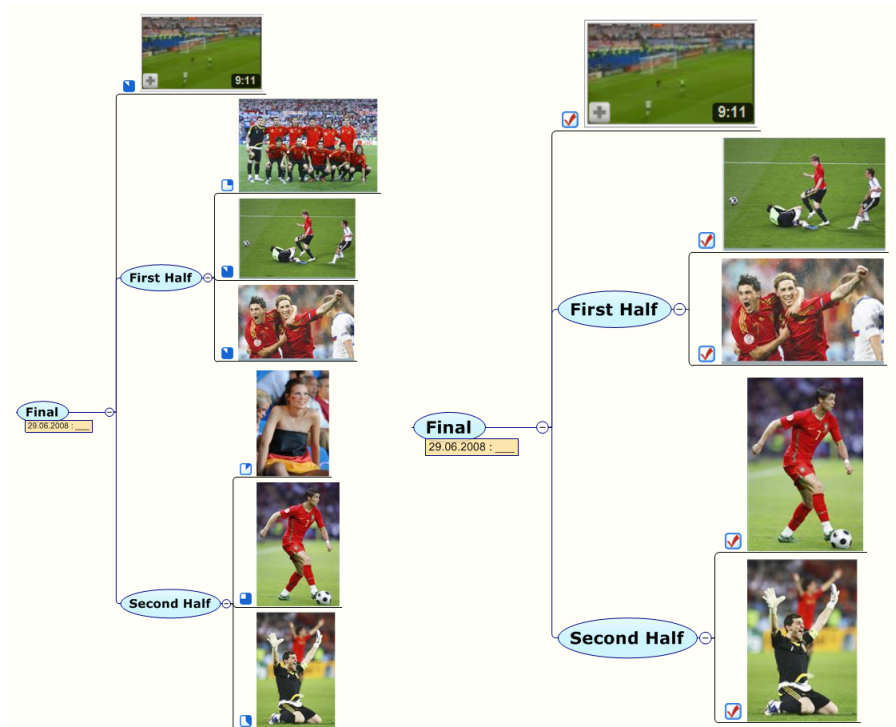


Fig. 4. Illustration of how GLOCAL can propose a classification of a user's new media (left) and how the user can second-guess the system (right).

A familiar-sounding example of this general approach is an author who creates a properly formatted submission to the SETN 2010 conference by taking a Word document with a submission to the SETN 2008 conference:

- If the structure of the author's new submission is exactly parallel to the structure of the old submission, all the author has to do is replace the original content with his own content. He may not have to think explicitly about the structure at all.
- Even if the structure of the old document is not quite right, the author can adjust it in an ad hoc way in the new document, without having to think in general terms about document structures. For example, he might add an appendix using the same format as for one of the normal sections of the paper.
- An intelligent system could support this type of activity by comparing the user's new document with other SETN 2008 (or similar) papers and perhaps suggesting improvements in the structure (e.g., a slightly different way of formatting a section that has the title "Appendix" and comes at the end of the paper).

In Figure 5, we assume that the user who wants to add media of her daughter's soccer tournament has already seen the event structure for Euro 2008 and has therefore decided to copy it as a starting point for the new tournament. She has recognized the need to simplify the structure somewhat and has renamed a couple of the subevents. For

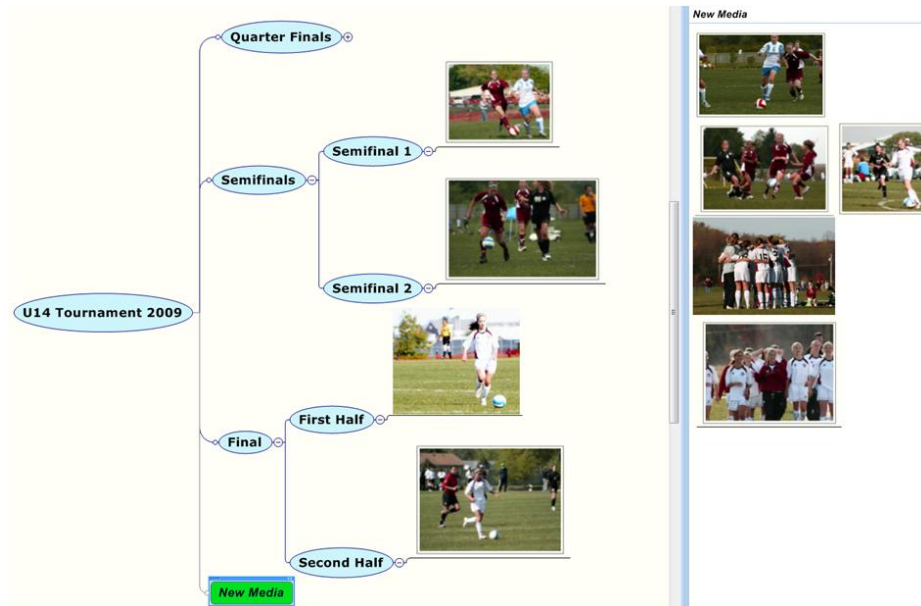


Fig. 5. Illustration of a situation in which a user has (a) created a new complex event instance, using an existing event instance as a starting point; and (b) inserted a small number of media into the new structure so as to enable the system to insert the other new media.

example, the youth soccer tournament does not have a distinction between a “Group Stage” and a “Knockout Stage”; it begins directly with the quarterfinals.

The figure shows the state of the system after the user has (as in the previous use case) uploaded her “new media”, which concern various games in the tournament, and assigned one medium to each leaf node in the hierarchy. Note that it is necessary for the user to do this initial work of placing some media in the appropriate places, since in this situation the system initially does not know any details about the subevents represented by the nodes and can therefore not perform an initial tentative categorization of new media, as it did in the previous use case.

The system now has some information about the times and places of the games, about the colors of the teams’ uniforms in each game, etc. Given this information, the system can guess at the classification of the remaining media, as before (the confidence levels are not shown in the figure).

But it is unlikely that all of the media will fit naturally into the structure that the user has just created, given that this structure was simply created ad hoc on the basis of a structure for another complex event. We must assume that there may be media that call for some adaptation of the event structure.

In our example, as shown in Figure 6, the system notes that the last two photos don’t seem to fit into any subevent. The system might conceivably ask the user to extend the event structure to create a slot for them, but most users would find this operation difficult.



Fig. 6. Illustration of how the system might suggest an improvement on the users structuring of the complex event, making use of existing representations of similar events.

So instead, the system examines the structures of other complex events (in this case: soccer tournaments) that have been created and used in the past. It notices that some of these events have included a “Celebration” event right after the end of the final game.

So it tentatively introduces this event node, putting the questionable media under it and offering an explanation of why the new subevent seems reasonable.

If the user doesn’t like the suggestion, she can ask the system to suggest other subevents in a similar way (or she can just delete the photos, if she can see that they are irrelevant).

6 Related Work

A great deal of research on support for photo annotation—mostly not involving indexing in terms of events—has yielded many ideas about effective combinations of

backend processing and interaction design (see, e.g., [5], [1], for individual contributions and [3] for a brief synthetic overview). Some of the work in this area also refers to indexing in terms of events. Some research (e.g., that of [2]) focuses on the technical aspects of event clustering. [4] likewise explore event clustering somewhat similar to the type of clustering assumed in the scenarios in this paper, also providing evidence for the viability of the sort of collaboration between user and system that is proposed here.

7 Conclusions and Next Steps

These scenarios and hypothetical examples illustrate how it may be possible and natural for untrained users to (a) benefit from an organization of media in terms of complex event structures and even (b) to create new event structures themselves, as a natural by-product of organizing their own media.

We are currently working on variants of these scenarios, which will then be presented to typical potential users, whose responses will presumably suggest desirable changes. The subsequent step will be the implementation of mockups that allow the interaction design to be tested.

These scenarios do make some strong assumptions about the capabilities of GLOBAL's backend processing, which is being developed in parallel in other parts of the GLOBAL project. Understanding of how the interaction can work helps to guide the development of the backend processing, and vice versa.

References

1. Barthelmess, P., Kaiser, E., McGee, D.: Toward content-aware multimodal tagging of personal photo collections. In: Proceedings of the Ninth International Conference on Multimodal Interfaces. pp. 122–125 (2007)
2. Cooper, M., Foote, J., Girgensohn, A., Wilcox, L.: Temporal event clustering for digital photo collections. *ACM Transactions on Multimedia Computing, Communications and Applications* 1(3), 269–288 (2005)
3. Hasan, T., Jameson, A.: Bridging the motivation gap for individual annotators: What can we learn from photo annotation systems? In: Proceedings of the First Workshop on Incentives for the Semantic Web at the 2008 International Semantic Web Conference. Karlsruhe, Germany (2008)
4. Suh, B., Bederson, B.B.: Semi-automatic photo annotation strategies using event based clustering and clothing based person recognition. *Interacting with Computers* 19, 524–544 (2007)
5. Tuffield, M.M., Harris, S., Dupplaw, D., Chakravarthy, A., Brewster, C., Gibbins, N., O'Hara, K., Ciravegna, F., Sleeman, D., Shadbolt, N., Wilks, Y.: Image annotation with Photocopain. In: Proceedings of the First International Workshop on Semantic Web Annotations for Multimedia, held at the World Wide Web Conference (2006)