Learning Café: a semantic multimedia collaborative platform for e-learning

Farès Belhadj, Vincent Boyer, Guylain Delmas, Myriam Lamolle, Chan Le Duc
LIASD-EA4383
University of Paris8
Saint Denis, France
{amsi, boyer, delmas, lamolle, leduc}@ai.univ-paris8.fr

Pierre-Antoine Champin, Christine Michel
LIRIS-CNRS
University of Lyon 1
Lyon, France
pierre-antoine.champin@liris.cnrs.fr, christine.michel@insa-lyon.fr

Abstract

The main purpose of our study is to design a new multimedia learning platform to help staffs to acquire new skills for their professional progression. In particular, this new platform is aimed at learners who do not appreciate conventional training methods, and who prefer new kinds of media, personalized to their own background. From a learner’s point of view, the platform provides semantic searches for media and proposes automatically the most appropriate video according to the user’s personalized pathway. That video can be rendered according to the user’s profile for enhancing specific information. In addition, the platform allows users to add new video resources, which will be automatically annotated, and to update the representation of learning domains. The present paper proposes a general design of the platform and the main ideas for addressing the raised research problems.

1. Introduction

In the craft industry or manual work, the expertise is developed by the acquisition of gestures or precise procedures. Professionals usually acquire hands-on experience and can benefit from being guided by an expert journeyman. They can also follow a formal training program mixing lectures and field practical work. The important feature of these learning programs is to show and explain gestures and job procedures.

The objective of our work is to build a training platform, named Learning Café, that is aimed at meeting those needs. We adopt a training method based on videos which precisely show gestures and working contexts. More specifically, the platform allows:

- to store and to qualify exactly videos in order to be exploited in various ways for learning purposes,
- to build personalized training courses.

The originality of our approach is to make the stages of storage, qualification and personalized access to be practicable by non-specialists such as learners or trainers. This facilitates the use of the system in various contexts and allows it to be updated continuously. From the user’s point of view, the platform is composed of two parts:

1. The interface for storage and indexation, relies mainly on a video annotation system. It allows the user to define the metadata required in the process of customization. These metadata include the annotation of key elements the video such as gestures and job procedure. The use of ontologies, both of the professional domain and the document structure, enable an accurate handling of those key elements by the system.

2. The visualization interface adapts the videos according to the user’s profile. The personalized pathway engine chooses the sequence of videos to propose to the user and performs graphical processing on the videos. This would make important elements of the videos more visible. Three main pathways are proposed: serious game, e-learning, direct consultation with keyword-based search.
Figure 1 presents a global view of the Learning Café platform and focuses on the process of personalized pedagogical pathways.

To illustrate our presentation, we use the same example along the paper. We consider a use-case focused on carpentry skills, for which a set of annotated videos and the domain ontology are available. In this context, we imagine that a user, a carpenter, wants to add a new video about safety at the customer's site. As shown in Figure 1, this new video is annotated to describe its pedagogical content (the arrow between "users" and "annotation extraction"). Note that the user can realize multiple annotations on the same video. Thus the ontology is dynamically revised and any new user-request on related topics will take into account this new video in the path generator and the expressive rendering engine.

The purpose of this paper is to present the general design principle of the platform. We first introduce the key elements of the design: ontologies, dynamic plot generation, video annotation and expressive rendering. The interactions between those elements are represented in Figure 1 and will be illustrated in the respective sections. We present in Section 6 the user-centred design process, based upon three stages of evaluation.

2 Ontologies and dynamic revision

The Learning Café platform needs to represent different ontologies to facilitate the generation of dynamic narrative scenarios in personalized pathways (see Figure 1). The first ontology, namely the profile ontology, represents the user's profile among which the training tracks. The second one, namely the domain ontology, concerns the learning domain of the company business. The third one, namely the video ontology, takes into account the annotated videos to classify them. To establish semantic relationships between these ontologies, we need to align them to obtain correspondences between entities which are present in the ontologies. For instance, there is a correspondence between two classes Video that appear in the videos and domain ontologies.

The profile ontology allows to store the videos known by a user thanks to the relationships commentedBy, seenBy, notedBy, etc., between the class Video and the class Learner. In the domain ontology, the class Video contains pedagogical resources which are defined in terms of the SCORM standard\(^1\). In the video ontology, the class Video contains technical descriptions with metadata such as duration, annotator and version. Each video illustrates a skill associated to competences in the domain ontology. A video may be a trimmed version of another one (trimmed- VersionOf relationships).

When a learner begins or follows a training, the Learning Café platform compares the chosen training and her profile to detect the level of corresponding capacities. To acquire a new competence \(c\) for a job \(j\), a learner \(l\) can send a query to the pathways generation module. When processing the query with the triple \((l, c, j)\), this module searches for relevant videos by using the knowledge inferred from the ontologies, the alignments, and the level of the learner \(l\) (beginner, advanced, etc.). The set of the videos that match the query can be pre-ordered according to the order of courses inferred from the domain ontology and the alignments. A graph (i.e. a set of personalized pathways) of videos, corresponding to a set of skills, is built from the set of videos, eventually taking into account the knowledge from the video ontology.

In the use-case introduced in Section 1, a learner wishes to watch videos focused on carpentry to acquire competences about safety (cf. Consultation in Figure 1). She should send a query to the pathways generation module (Figure 1) and receive personalized pathways. The parameters of the query provide the information such as the learner's competences, the carpentry job, safety, etc., which are encapsulated in the triple \((l, c, j)\) as mentioned above.

Semantics-based applications usually encapsulate a set of ontologies, which represent the knowledge involved in different data sources. Some of them evolve constantly since the data sources, the applications environment and the users' needs change over time. Among the first works related to dynamic revision of a knowledge base, the AGM framework [1] introduced a set of logic postulates to be

\(^1\)http://www.adlnet.org/scorm
satisfied when defining revision operations. Among these logic constraints, knowledge retraction and minimality of changes are the most problematic. Recently, Ji et al. [9] have presented the RaDON system, which can check consistency of a simple ontology and a network of ontologies when revising knowledge. This system is based on a "relaxed" notion of ontology consistency, and only supports the revision of inexpressive ontologies.

In the context of the Learning Café platform, the ontologies may change according to the business and new skills of the enterprises. In particular, the video ontology should be integrated with new annotated videos from diverse sources. The annotations may imply an evolution of the ontology models since new concepts and/or relationships or new contexts of use should be taken into account.

In the use-case introduced in Section 1, the dynamic revision module is triggered when the new annotated video about carpentry is added. This module extracts new terms from the annotations of this video to enrich the domain ontology. If the updated domain ontology becomes inconsistent, the dynamic revision module either asks for an intervention of experts or manages to ensure consistency of the domain ontology by deleting some incompatible knowledge.

3 Dynamic plot generation

Besides generating for every learner a personalized learning course, the Learning Café platform must be able to develop for each learner a course by taking into account feedback from the other learners. The Learning Café platform can use the domain, video and profile ontologies to not only generate but also recompute every learner’s course. The platform will achieve this by using dynamic plot generation methods. At the end of every video, the platform will take into account the learner’s feedback and will determine the most relevant videos to propose to him according to his personal feedback, the evaluation of videos by the other users and his progress in the learning sequence. The field of interactive narrative and plot generation has numerous results, the more noticeable are the narrative control of MIMESIS [19], the elaborated staging of Façade [13], the Interactive Drama Architecture from the University of Michigan [11], the work from the Intelligent Virtual Environments on hierarchical task networks [5] and the integration of user emotional feedback [6], then the theory of emergent narrative [3] or project ID Tension [15]. However the known approaches work only on the basis of a single user and of a globally homogeneous public. Their adaptation to an e-learning context thus raises questions about performances (to manage the load of a high number of simultaneous users), heterogeneity of the targeted public (which prevents certain simplifications of the proposed scenarios and requires a bigger capacity of adaptation from of the engine) and multi-users management (to propose different scenarios to several different users but presenting points of synchronization for the collective activities, to ensure a certain level of coherence in the proposed courses and to allow the users to exchange about their experiences). So, the Learning Café platform is going to integrate a dynamic plot generation engine capable of integrating not only the three ontologies defined previously, but also of considering courses followed by other users. For example the activities involving several users will be proposed when a sufficient number of learners can take place there, to make sure that a given learner is not blocked by the absence of the others. However, this objective to synchronize various users can generate conflicts between the various proposed courses. Our proposition’s stake is to find a method allowing these potential conflicts to be arbitrated in order to guarantee that no personal course is itself completely blocked by other users’ ones, and repercussions of a peculiar user’s course have no great negative impact on the other possible plots. To achieve it, we propose a software architecture implying 2 levels:

a plot generator: specific to each user, it receives the user’s input as reported by the platform, draws up plot propositions, reports them to the universe manager (see below) for validation and delivers the final instructions to the platform for execution.

a universe manager: it receives the plot propositions from the various plot generators and makes arbitrations when it detects potential problems (contradictions, blocked resources, incoherence).

To achieve its goal, the universe manager will rely on the domain and video ontologies. These ontologies enable it to follow up the modifications brought to the learning environment by every user’s plot and to calculate for every proposition how it can affect the other plots. The plot generator will dispose of a specific version of the ontologies focused on the needs of the concerned user and updated by the universe manager. This architecture should allow separating the activities of course generation and of management of the learning environment. The universe manager can therefore arbitrate between the various rival plot propositions by keeping a neutral position. Moreover this division allows distributing the calculations between user’s clients and a central server. Each various plot generators can work locally on every user’s computer, and the universe manager may take place on a cloud server.

4 Video Annotation

Video annotation in the Learning Café platform is done according to the principles and models developed for the
Advene project [2] (see Figure 2). Advene separates the annotation structure into three parts:

- An **annotation** in Advene is a piece of data (the content of the annotation) attached to a fragment of a video. There is no limitation about how the data has been produced (either manually coded by an analyst, or automatically extracted from the video signal) or will be rendered.

- **Annotation schemas** aim at capturing a way of annotating videos, by defining a set of annotation **types**. Each annotation type specifies the kind of data that those annotations may contain (simple text, image, structured data, etc...), and provides an informal semantics for them. For example, one type can be used to convey safety advices (textual content), while another type can point to dangerous situations in the video (structured data describing a zone in the image).

- **Views** are a mean to specify how annotations of a schema are to be rendered, either while playing the video or in a separate document. For example, the “safety advice” annotations can be rendered as subtitles while playing the video, or arranged in a document to generate a summary of the video.

In the **Learning Café** platform, video annotations will be produced at different stages, and used for different purposes. Some annotations are created at production time, either to describe the pedagogical content of the video using the domain ontology (see Section 2), or to prepare the expressive rendering (see Section 5). Other annotations will be created at runtime, either by explicit actions of the users (e.g. ranking or sharing the video with friends) or by implicit actions (e.g. stopping the video before the end). We have defined a schema for the former kind of annotation, and are currently working on extending it for the latter kind.

Views, on the other hand, are the core research question in the **Learning Café** project, as which videos each user will watch, and how they will be rendered, will be dynamically determined by the annotations but also by the user’s own profile, as described above in Section 3.

While Advene has its own format for representing all the elements described above, we will need to export (at least a part of) the annotation structure to RDF [12] in order to integrate it with the various ontologies described above. This export will take advantage of existing standards for identifying video fragments [16] and expressing standard video metadata [10]; but of course, dedicated vocabularies may have to be proposed as well.

---

2 [http://advene.org/cinelab](http://advene.org/cinelab)

---

5 Expressive rendering

Images and Videos convey information to the viewer. These information can be enhanced through an expressive rendering process for a more efficient visual communication. In the **Learning Café** platform, we aim to automatically produce expressive videos according to the topic and user’s profile.

Expressive rendering (previously named non-photorealistic rendering) focuses on artistic and illustrative techniques. Illustrative techniques make efficient images noticeably by omitting extraneous details, clarifying shapes, focusing attention, illustrating ideas, conveying emotions.

In the **Learning Café** project, we focus on both abstraction techniques [7] used to clarify the image structure and stylization techniques [8] that convey information and emotions. Objects, shapes or areas are emphasized or abstracted and the viewer will focus its attention on important parts of the video. Stylizations will be used to convey emotions. Annotations (see section 4) are used to produce abstractions while stylization is deduced form the user’s profile (see section 3). Thus considering a topic, the selected video, its annotations and the user’s profile, the rendering process will produce a specific video through abstraction and stylization techniques. The expressive videos will always be rendered on the fly in an interactive time and will be never stored on the **Learning Café** platform.

As the **Learning Café** project deals only with videos, successive images must be abstracted and stylized while preserving a good temporal coherence and avoiding popping effects. Recently, Benard et al. [4] have summarized the trade-offs made by various solutions for the temporal coherency problem. They have noticed the issue conflicting nature and have stated that the solution is necessarily a compromise. Currently, for the stylization, there are only solutions for static images and no solution exists for videos (i.e
the temporal coherence remains a problem). Concerning the abstraction techniques, even if our model will be inspired from related work like [18], as a preliminary work, we have developed a solution providing edge detections mixed with an abstraction blur effect applied on specific areas annotated using the Advene platform. This preliminary solution is presented on figure 3 and matches the use-case given in the introduction. We suppose that annotations about safety and stair building have been attached to the video by the user and we propose the corresponding abstractions. This solution will be used in the first evaluation campaign (see section 6). On the produced images, the mouse pointer is still presented to illustrate the annotated area.

Thus the main expressive rendering challenges will consist in producing efficient video communication through abstraction techniques that takes into account Advene annotations and produce temporally coherent stylizations.

6 Evaluations

The process of use is decisive in the case of the Learning Café project. Indeed, it is required to make the system evolve according to the information gained through interaction. The uses are also needed to evaluate the relevance of the research proposal, specifically about personalization processes. That’s why we suggest to use a user-centred design methodology and so organize campaigns of evaluation in three key stages:

- In the beginning of the project, we suggest estimating the perceived utility of Learning Café for the professionals and their preliminary attitudes regarding the use of videos or learning games found on the web to learn gestures or processes, as well as their habit to use annotations or comments on social networks. This evaluation is made in the form of a questionnaire.

- During the design process, we suggest testing the usability of specific parts of the platform: the annotation interactions and the video consultation interfaces for information or training purposes. We propose to observe the way users use the platform (by direct observation and analysis of traces) and their satisfaction statement (though a survey). The tested elements are the criteria chosen for annotation, relevance of information presented in the video, tasks done with videos and functions chosen to perform the browsing process. Results of the studies are used to refine the design of the ontology (analysis of the comprehensiveness and precision of the context description), the personalization rules used to build the dynamic scenario, the stylistic effects applied to videos and the ergonomics of the interfaces. These evaluations are also useful to estimate the quality of the produced contents. Indeed, the annotation process is partly done by the designers (for the automatic annotation or the reuse of previous ones) and partly done by the users (manual annotation) [14]. These processes produce numerous errors. Indeed, during the evaluation of its system, [20] identified computational errors in the calculations of interpolation due to the movements of objects or to the vibrations of the image if the camera moves. He also identified numerous human errors in annotation qualification relative to objects or events.

- At the end of the design process, we suggest to do a real and global use test with a sample of twenty professionals and their managers during a whole training course. The first objective is to measure the learning process (in terms of skills and new behaviour gained) and the user experience (UX) in particular the satisfaction and the feeling of evolution. The second objective is to measure the efficiency [17] of each research proposal for learning and UX building. Indeed we want to distinguish the effect of annotation interaction, the watching of video according to the scenarios and the watching of video according to the browsing process. In order to do that we plan to test various scenarios by the way of comparative experimentations with a test group using the whole platform, a simple group using the platform with classical video (i.e. without annotation and effect) and a control group using a classical learning process. The third objective is to build a model of continuous design process. The uses produce various evolutions on the platform and we want to pro-
pose a model for the evolution of the set of annotation (proposed by the user, automatically extracted by the system or chosen for each user), the evolution of the ontologies (new concepts identified, new links made) and of the personalization rules.

7. Conclusion

We have presented the general design of the Learning Café platform dealing with the raised research problems. This 3 years project has started in May 2013 and is funded by the Ministry of Economy, Industry and Employment of the French Government. In this project, each research problem (i.e. ontologies, dynamic plot generation, video annotation and expressive rendering) will be investigated. Moreover the consortium (LIASD, LIRIS, Omedo, Xsalto, La Talemerie, INBP, Maison & Services, Mandarine Codi) also comprises end-users of the system such as home service companies, and bakeries. Employees of those companies should acquire particular skills for their professional progression and these companies are particularly motivated to provide any media and organize campaigns of evaluation.

Acknowledgment The preliminary expressive rendering solution presented above was developed by Grafeet S.A.S.

References