Citation for published version


DOI
https://doi.org/10.1142/S0218194010004839

Document Version

This is the Accepted Manuscript version.
The version in the Universitat Oberta de Catalunya institutional repository, O2 may differ from the final published version.

Copyright and Reuse

This manuscript version is made available under the terms of the Creative Commons Attribution Non Commercial No Derivatives licence (CC-BY-NC-ND)
http://creativecommons.org/licenses/by-nc-nd/3.0/es/, which permits others to download it and share it with others as long as they credit you, but they can’t change it in any way or use them commercially.

Enquiries

If you believe this document infringes copyright, please contact the Research Team at: repositori@uoc.edu
This study aims to explore two crucial aspects of collaborative work and learning: on the one hand, the importance of enabling collaborative learning applications to capture and structure the information generated by group activity and, on the other hand, to extract the relevant knowledge in order to provide learners and tutors with efficient awareness, feedback and support as regards group performance and collaboration. To this end, in this paper we first propose a conceptual model for data analysis and management that identifies and classifies the many kinds of indicators that describe collaboration and learning into high-level aspects of collaboration. Then, we provide a computational platform that, at a first step, collects and classifies both the event information generated asynchronously from the users’ actions and the labeled dialogues from the synchronous collaboration according to these indicators. This information is then analyzed in next steps to eventually extract and present to participants the relevant knowledge about the collaboration. The ultimate aim of this platform is to efficiently embed information and knowledge into collaborative learning applications. We eventually suggest a generalization of our approach to be used in diverse collaborative learning situations and domains.

Keywords: Collaborative Learning; Knowledge Management, Ontology, Awareness, Feedback, Scaffolding, Group Monitoring.
1. Introduction

Computer-Supported Collaborative Learning (CSCL) is a paradigm for research in educational technology that focuses on the use of Information and Communications Technology as a mediation tool within collaborative methods of learning [1]. One key issue when developing CSCL applications is interaction data analysis, a core function for the support of coaching and evaluation of the collaborative learning process. CSCL applications are characterized by a high degree of user-user and user-system interaction and hence are generating a huge amount of quantitative information (log files) from both synchronous and asynchronous collaboration.

The constant and fast processing of this quantitative data source collected as well as their systematic analysis based on principled indicators (variables) enable the measurement of the type and the degree of group members’ participation [2]. The knowledge extracted by this process can then be used to facilitate a continuous monitoring of the learning activity, providing group members with appropriate support, as well as awareness [3] and feedback [4] about what is happening during collaboration. The presentation of this knowledge to the interested actors may positively impact on participant’s motivation, emotional state and problem-solving abilities and as a result enhance on-line collaborative learning [3].

In addition, qualitative information is collected from ad hoc questionnaires which are regularly filled out by group members, reporting human and behavioral aspects of collaboration as well as evaluating the collaborative learning experience. Participants qualify their own emotional and motivational state within the learning group as well as evaluate the participation and learning activities of their peers. The aim of this qualitative approach is to provide both a deeper understanding of collaboration and a more objective assessment of individual and group activity.

The ultimate aim of our work is to extract relevant knowledge of the collaboration process from all possible sources. Note that in this context information refers to quantitative and qualitative data generated by the learning group whereas knowledge refers to the result of the treatment of this information through analysis techniques and interpretation. The development of a clear and well-structured conceptual model can facilitate the building of a portable, general and reusable collaborative learning representation and inference of knowledge about each collaborative process [5]. The whole approach is based on our experience in the real context of learning of the UOC1.

In order to achieve these goals, in Section 2, we extended a previous conceptual model [6] for data analysis and management in order to identify and classify the many kinds of indicators that describe collaboration and learning into the above-mentioned potential aspects of collaboration. Then, in Section 3, this conceptual approach is translated into a computational model that constitutes a generic platform for the systematic construction of CSCL applications with enriched capabilities for knowledge

---

1 The Open University of Catalonia (UOC) offers full distance education through the Internet since 1994. About 54,000 students and 2,500 lecturers and tutors are involved in more than 1200 on-line official courses from about 30 official degrees and other PhD and post-graduate programs. The UOC is found at http://www.uoc.edu.
management and group scaffolding. Last Section summarizes the paper and outlines ongoing efforts by pointing out the experiences performed so far from real CSCL practices supported by our approach as well as suggesting a generalization of the results of this research to be used in other contexts of collaborative learning practices.

2. A Conceptual Model for Managing Group Activity Interaction

The model we propose in this paper (see Figure 1) is extended from a previous conceptual model [6] so as to further support synchronous communication as well as collect qualitative data from the collaborative learning experience. The whole approach aims at modeling different aspects of interaction and thus at helping all the actors involved understand the outcomes of the synchronous and asynchronous collaborative learning process. We therefore base the success of CSCL applications on the capability of such applications to embed information and knowledge extracted from group activity interaction and use it to achieve a more effective group monitoring. The essential issue here is how to manage the information from real, long-term, complex collaborative problem solving situations in order to extract relevant knowledge from group activity with the aim of providing learners with efficient awareness and feedback as regards individual and group performance and assessment as well as enabling the instructor to both analyze group interaction effectively and provide an adequate support when needed.

Figure 1. An UML excerpt of the proposed conceptual model.

Quite a few ontologies [5] and related standards [7] concerning the representation of CSCL have been defined so far. Representative approaches include [5] that use a combination of a general domain ontology describing the common semantics needed for the implementation of a collaboration environment with several domain ontologies that
are used to provide a framework for end-user tools. Barros et al [8] propose to use the actions performed in the collaborative learning system so as to build a high-level representation of the process of collection and analysis of the interaction data. In [9] a theory-oriented interaction analysis approach based on theories of collaborative learning is provided. However, the social processes happening behind real collaborative learning practices are very complex and subjective and thus they fall far from a holistic view proposed by standards and ontologies [10]. We believe that each and every collaborative learning setting needs a particular conceptualization, with no prior consideration for the implicated actors, and thus predefined concepts and categories should be used to model the analysis of the interaction data generated. This inductive approach sharply contrasts with the deductive methodology inherently used by ontologies, which propose a unique view that frames the standardization of any collaborative learning process.

Following an inductive methodology, we classify group activity information generated in our real context of learning into three generic categories of activity [2]: the performance of the task (the outcome of collaboration or the members’ contributing behavior to the task), the functioning of the group (the management and organizational processes underlying the collaborative learning activities, such as participation behavior, role playing, etc.), and individual and group scaffolding (social support and task- or group functioning-oriented help). Since that model was initially built for asynchronous collaboration, we extend it to cover the synchronous case as well.

<table>
<thead>
<tr>
<th>Skills</th>
<th>Sub-skills (Learning outcome contribution)</th>
<th>Asynchronous actions (A)</th>
<th>Synchronous communicative acts (S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Basic active learning skills</td>
<td>Information generation</td>
<td>Create doc/note (A)</td>
<td>Describe / explain (S)</td>
</tr>
<tr>
<td>Supporting active learning</td>
<td>Information refinement</td>
<td>Edit doc (A)</td>
<td>Adj ust (S)</td>
</tr>
<tr>
<td>Supporting active learning</td>
<td>Information elaboration</td>
<td>Version / Replace doc (A)</td>
<td>Elaborate (S)</td>
</tr>
<tr>
<td>Supporting active learning</td>
<td>Information revision</td>
<td>Revise / Branch doc (A)</td>
<td>Revise (S)</td>
</tr>
<tr>
<td>Supporting active learning</td>
<td>Information reinforcement</td>
<td>Create_Noteboard doc / URL / Notes (A)</td>
<td>Extend (S)</td>
</tr>
</tbody>
</table>

Next, we briefly describe each of these three categories and their associated skills (see [6] and Figure 1). We employ a similar terminology to the one used in the Basic Support for Cooperative Work (BSCW) system [11] to refer to the actions that can be carried out...
in an asynchronous groupware platform. However, they are general enough to represent all the typical and basic asynchronous interaction encountered in the different programs and studies in our real learning context of the UOC. On the other hand, based on [2] we use specific terminology for labeling the dialogues generated in collaborative synchronous environments.

2.1. Collaborative learning outcome (or task performance)

Table 1 shows the mid- and low-level indicators in the form of the skills and sub-skills that should characterize the students who participate in a learning collaborative situation in order to achieve effective group and individual performance of the task and thus obtain a successful learning outcome. To measure each indicator (or skill), we associate it with both the actions that students perform in an asynchronous (A) environment and the type of dialogues carried out synchronously (S).

2.2. Group functioning

Table 2 shows the mid- and low-level indicators in the form of skills and sub-skills that students should exhibit in order to enhance participation, accomplish well-balanced contributions, promote better communication and coordination, as well as adequate workload distribution, task management and workspace organization. The aim is to achieve an effective group interaction and functioning in a collaborative learning situation. To measure each indicator, we associate it with specific student action and contribution types which best describe each skill to be accomplished.

<table>
<thead>
<tr>
<th>Skills</th>
<th>Sub-skills (Group functioning contribution)</th>
<th>Asynchronous actions (A)</th>
<th>Synchronous communicative acts (S)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active participation and peer</td>
<td>Participation in managing information</td>
<td>Create Event, Change Event, Read Event (A)</td>
<td>Take-initiative, Provide-info, Share-info, Request/Suggest-action, Listen (S)</td>
</tr>
<tr>
<td>Social grounding skills</td>
<td>Well-balanced contributions, adequate reaction attitudes, and role playing</td>
<td>Create Event, Change Event, Read Event, Move Event (A)</td>
<td>Provide-acknowledgment/answer/solution, Assess, Give/Take-turn, Perform-role (S)</td>
</tr>
<tr>
<td>Task processing skills</td>
<td>Task planning/distribution</td>
<td>Create/Link Appointment ;</td>
<td>Create/ChangeAccess WSCalendar (A)</td>
</tr>
<tr>
<td>Task (and knowledge) management</td>
<td>Work load distribution</td>
<td>Create Folder ;</td>
<td>Create Notes (as a contribution in a bulletin board) (A)</td>
</tr>
<tr>
<td>Work load distribution</td>
<td></td>
<td>Build-workspace, Distribute-workload (S)</td>
<td></td>
</tr>
</tbody>
</table>
2.3. **Scaffolding**

Table 3 shows the different types of social support and help services [12] that have been identified and accounted for in our model. The participants’ actions and contributions aiming at getting or providing help are classified and measured according to whether they refer to the task or group functioning.

Table 3. Indicators that model scaffolding (for both asynchronous and synchronous collaboration).

<table>
<thead>
<tr>
<th>Social support</th>
<th>Help Services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Members’ commitment toward collaboration, joint learning and accomplishment of the common group goal</td>
<td>Help is timely</td>
</tr>
<tr>
<td>Level of peer involvement and their influential contribution to the involvement of the others</td>
<td>Help is relevant to the student’s needs</td>
</tr>
<tr>
<td>Members’ contribution to the achievement of mutual trust</td>
<td>Help is qualitative</td>
</tr>
<tr>
<td>Members’ motivational and emotional support to their peers</td>
<td>Help is understood by the student</td>
</tr>
<tr>
<td>Participation and contribution to conflict resolution</td>
<td>Help can readily be applied by the student</td>
</tr>
</tbody>
</table>
Qualitative information about group functioning and scaffolding is also extracted by specifically designed structured and non-structured questionnaires which are filled by group members at the end of each collaborative problem-solving phase. Structured questionnaires provide a predefined set of answers to choose and as a result can be collected and processed by computers whereas non-structured questionnaires present a high degree of informality and thus need to be processed and interpreted manually. Table 4 shows a generic questionnaire scheme which is eventually elaborated and adapted to the particular problem-solving situation.

Finally, we consider group well-being [13] as a transversal function that attempts to incorporate aspects of human virtues and behavioral processes such as members’ emotional state, self-criticism and motivation into the analysis and interpretation of collaboration. This function adds another specific qualitative layer into the analysis process. Indeed, taking this function into account, we add a new dimension on group monitoring and decision taking since we are not based exclusively on the rational results extracted from quantitative (or other qualitative) data analyses. In our model, information about group well-being function is collected by qualitative data generated in the form of report results during the collaboration. The interpretation of these results allows the tutor to understand and evaluate the learning process more objectively as well as to identify and correct misleading behavior of the collaborative partners. As a result, the tutor (or the group coordinator) is able to provide adequate feedback that may increase participants’ motivation and emotional state [3] and as a result increase both the quality and the quantity of group activity regarding each of the three categories described above.

In order to extract reliable qualitative information about group’s well-being function, students have to fill out both structured and non-structured ad hoc questionnaires, as the ones mentioned above. Table 5 presents a generic questionnaire scheme that aims to extract information about group participants’ motivation, emotional state as well as self, peer and group activity evaluation.
Table 5. A generic questionnaire scheme about the group’s well-being function.

<table>
<thead>
<tr>
<th>Question</th>
<th>Response Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indicate your own motivational/emotional state at this stage of work</td>
<td>Scale 0-5</td>
</tr>
<tr>
<td>According to your knowledge, indicate the motivational/emotional state of your peers and the group as a whole (0 – 5).</td>
<td></td>
</tr>
<tr>
<td>Expose the reasons that explain your motivational/emotional state (Text).</td>
<td></td>
</tr>
<tr>
<td>As far as you know, describe the motivational/emotional state of your peers and the group as a whole (Text).</td>
<td></td>
</tr>
<tr>
<td>Indicate your benefits from the online collaborative learning experience so far (Text).</td>
<td></td>
</tr>
<tr>
<td>Indicate how optimist you feel as regards the achievement of a successful collaboration at the end of the experience (0 – 5).</td>
<td></td>
</tr>
<tr>
<td>Indicate your expectations at this stage of group activity (Text).</td>
<td></td>
</tr>
<tr>
<td>Have they been fulfilled? (YES, NO, NR/NS).</td>
<td></td>
</tr>
</tbody>
</table>

The following problems to be faced are: (i) how to process the large amount of both quantitative and qualitative information collected during group activity efficiently in order to facilitate its later analysis and make the extracted knowledge available to the participants even in real time; (ii) how information should be analyzed and what kind of knowledge should be extracted to be presented to the participants in order to provide the best possible support and monitoring of their learning and instructional processes. Next section proposes a solution to these problems by providing an efficient and reusable computational approach that enables the embedding of the collected information and the extracted knowledge into a CSCL application.

3. **A Computational Model for CSCL Applications**

A generic, robust, reusable component-based Collaborative Learning Purpose Library (CLPL) [14] was developed as a computational model so as to enable a complete and effective reutilization of its generic components for the construction of specific CSCL applications. This platform implements the conceptual model of information management described in Sect. 2.

The CLPL is made up of five components related to user management, administration, security, knowledge management, and functionality mapping the essential needs in which any CSCL application is involved. Special attention has been paid to addresses the complex issues of data analysis and management identified in the previous section. This is mainly performed by two components, namely **CSCL Knowledge Management** and **CSCL Functionality** components, which form the core of the CLPL. Due to their importance, these two components are briefly described below (a detailed description is found in [14]).
3.1 CSCL Knowledge Management Component

In developing the CLPL we paid special attention to event analysis and management forming the three first stages of a process of transforming this information into knowledge as described in [6] (see also Figure 2). To this end, a generic log file is provided as a key entity that collects and classifies all the action events generated during group activity in a certain workspace over a given period of time, and constitutes a source of information that will be later processed by statistic techniques. Classification of event information is based on a complete and tight hierarchy of events (Fig. 1) based on the mentioned three types of collaborative activity proposed in Sect. 2.

Given the generated events which have been previously collected, classified and stored in log files, this component also performs the statistical analysis event information as well as the management and maintenance of the knowledge extracted by that analysis. To this end, a statistics abstraction is provided which takes into account both the information source stored in the database and the associated generic criteria that guide the performance of the desired quantitative analysis of individual and group activity. In
addition, a data structure of generic and parameterized criteria was also designed to address the most usual requests for information in CSCL environments (e.g. "How many users accessed the system during a period of time?", "Which users read a document?") thus making it possible to reuse them in as many statistics as possible.

The ultimate aim of this component is to define a bottom-up analysis approach that analyses the user events in order to decode the specific actions of the users describing their interaction during the collaboration activities. The analysis aims at identifying those sequences of actions that can be used to determine typical patterns of interactions.

### 3.2 CSCL Functionality Component

The final objective of this component is to provide functional support to CSCL applications in terms of group organization, resource sharing and user interaction. Moreover, this component implements the last stage of the process of embedding information and knowledge into CSCL applications (see Fig. 2) by presenting the knowledge generated to users in terms of immediate awareness and constant feedback of what is going on in the system.

This component distinguishes different levels that dictate how the acquired knowledge is to be presented, namely awareness, feedback, assessment and scaffolding (or guiding) levels [2]. At awareness and feedback levels, the aim is to inform participants about what is going on in their shared workspace, providing information about their own actions or the actions of their peers, or presenting a view of the group interaction, behavior and performance. At assessment level, this component provides data and elements to assess the collaborative activity, so the indicators used are associated with specific weights that measure the significance of each indicator in the assessment process. Finally, at scaffolding level, this component produces information aiming at guiding, orienting and supporting students in their activity.

In order to provide this information, certain key entities are defined in this component, such as resource state, user status and group memory [13]. Each of these abstractions acts as a vehicle so that the knowledge acquired can be classified and presented to users in the correct form depending on the type of activity involved. For instance, in resource sharing (e.g. a multi-user editor session), participants are continuously modifying the state of the shared application (e.g. writing a new text comment, deleting somebody else's sketch, etc.) and thus the current application state has to be continuously propagated to the users as a news warning signal. Furthermore, as a consequence of the complex knowledge provided to participants (e.g. group's member relative and absolute amount of contributions) this component defines certain generic entities such as history, pool and diagram and functions such as sorting. Based on these abstractions it is possible to dynamically gather and store great amounts of history data and statistical results from the group activity in order to constantly update and present them to participants in the appropriate diagrammatic form.

To sum up, the CLPL platform reflects and describes task performance, individual and group behavior, interaction dynamics, members' relationships and group support as
A Conceptual Model for Enhancing Knowledge Management in Online Collaborative Learning

accurately as possible. Furthermore, the genericity, robustness and reusability provided by this platform can be used for the systematic construction of CSCL applications endowed with enriched capabilities for providing more efficient knowledge management and scaffolding and group monitoring.

4. Conclusions and Ongoing Research

In this paper, we have discussed an approach for transforming information generated from different sources of learning group activity into useful knowledge in an efficient manner for individual and group awareness, feedback, monitoring and scaffolding. The aim is to enable group members to become aware of their own progress and that of their peers in performing a learning exercise, as well as of the extent to which other members are participating in the collaborative process as this influences their decision making. In addition, this approach provides tutors with information about students' problem-solving behavior, group processing, and performance analysis for assessment and guiding purposes.

We plan to incorporate the innovative research ideas presented into the CLPL platform and thus add further support for synchronous communication as well as qualitative extraction of knowledge. The acquired knowledge will then be used to improve the collaborative discussion processes happening in the real learning context of our virtual university. For validation purposes, we will use the updated version of the CLPL to enhance our successful prototypes of discussion forums [15], [16] built upon this platform. The promising results obtained so far by using these prototypes encourage us to keep working in this direction.

Instead of mapping our approach to data models of particular LMS (such as Moodle), we plan to integrate our approach to programmatic interface specifications, such as OKI-OSID (see http://www.okiproject.org/sites/oki-repository). As a result, the integration will be more generic and will not become obsolete after new versions of particular LMS. As a first step to this direction, we are working on the generalization of our approach by providing an integrated ontology that specifies the presented conceptual model refined with new information. We expect this initiative to provide a basis so that other collaborative learning domains can take advantage of our efforts to model both their specific needs and the way in which information is obtained in their learning contexts. The current version of our conceptual model written in OWL (the Web Ontology Language) [10] can be downloaded from http://clpl.uoc.edu/clpl_ontology.owl

Acknowledgments

This work has been partially supported by the European Commission under the Collaborative Project ALICE ”Adaptive Learning via Intuitive/Interactive, Collaborative and Emotional System”, VII Framework Programme, Theme ICT-2009.4.2 (Technology-Enhanced Learning), Grant Agreement n. 257639. Fatos Xhafa’s research work completed at Birkbeck, University of London, on Leave from Technical University of Catalonia (Barcelona, Spain). His research is supported by a grant from the General Secretariat of Universities of the Ministry of Education, Spain.
References


