The Data Flow Problem in Learning Design: A Case Study

Luis Palomino-Ramírez, Miguel L. Bote-Lorenzo, Juan I. Asensio-Pérez, Yannis A. Dimitriadis

School of Telecommunication Engineering, University of Valladolid, lpalomin@ulises.tel.uva.es, migbot@tel.uva.es, juaase@tel.uva.es, yannis@tel.uva.es

Luis de la Fuente-Valentín

Telematics Engineering Department, University Carlos III of Madrid, lfuente@it.uc3m.es

Abstract

A teaching-learning process formalized through the IMS-Learning Design specification (IMS-LD) comprises a sequence of learning activities (learning flow) as well as a sequence of outcomes which are the result of using tools and services in order to support the learning activities (data flow). Although the automation of the learning flow has been achieved in IMS-LD, the automation of the data flow between tools is still an open issue. Nevertheless, no real case studies have been reported in the literature, that illustrate the existence and relevance of the problem. In this paper an authentic case study is analyzed and several findings indicate that in collaborative learning the data flow management is error-prone for the students; the data flow specification is error-prone for the course designer; as well as a reusability issue in data flow specification has been identified. All of them are relevant issues since the two former may potentially affect the students' learning objectives while the latter affects the reusability of the whole unit of learning (UoL). In order to address these issues while keeping interoperability with IMS-LD, we introduce a solution that separate the data flow specification from the learningflow one.

Keywords

Learning Design, IMS-LD, data flow, reusability, case study.

Introduction

IMS-LD (IMS 2003) has greatly contributed to the formalization of teaching-learning processes via the standardization of the Learning Design (LD) approach. A learning design defines a flow of activities where students and teachers play roles while using tools and services with the aim of accomplishing their learning goals. Through a learning design, different actors do not have to be concerned much about the management of learning flow and data flow within a e-course: this should be performed automatically by the learning management system (LMS) (Koper and Miao 2007).

Nevertheless, the formalization of teaching-learning processes through IMS-LD has not been entirely accomplished. Although the automation of the learning flow has been successfully achieved (Hernández-Leo et al. 2005); however the automation of the data flow between tools is still an open issue in IMS-LD (Koper and Miao 2007) (Peter and Vantroys 2005) (Miao et al. 2005) (Wilson 2005) (Dalziel 2006) (Palomino-Ramírez et al. 2007). However, no case studies have been reported for any of the former authors in order to show with real data the existence and relevance of the problem. In (Peter and Vantroys 2005), (Wilson 2005) and (Dalziel 2006) the problem is simply enunciated and in (Miao et al. 2005) a case study is introduced without providing any real data. Finally in (Palomino-Ramírez et al. 2007) an illustrative example is presented, which is inspired by a real case but it concerns only fictitious data.

The aim of this paper is to present actual evidence of the IMS-LD data flow problem based on the analysis of an authentic case study in order to show why the automation of the data flow is a relevant

 ISBN No: 978-1-86220-206-1

issue in LD. Since the case study was not design for this purpose, real data in order to validate its significance and relevance to the problem should be also presented. The analysis of the experience will be carried out using data gathered from different sources: questionnaires, e-mail exchange among participants through the experience and the UoL used in the case study. A mixed evaluation method (Martínez-Monés et al. 2003) will be used in order to support our findings. By combining the expected findings and literature, it should be possible to identify distinct dimensions of the problem in order to define the necessary requirements for addressing it. Finally, it would be interestingly to know whether the composition-based approach proposed by us in (Palomino-Ramírez et al. 2007) in order to tackle the data flow problem could be used to address all the problem's dimensions that can be identified in this case study.

The rest of the paper is structured as follows. In next section a review of the IMS-LD data flow problem is introduced. Then, the actual case study is described, including the evaluation methodology. Next, the emerged findings through the case study analysis as well as the identified dimensions of the problem are presented. Then, a potential solution for addressing the problem is introduced. Finally, we summarize our conclusions and future work.

Review of the IMS-LD Data Flow Problem

Data flow in learning design refers to the sequence of artifacts which are outcomes that are produced and consumed by tools or services in the learning activities. Unfortunately artifacts are not part of the IMS-LD metamodel, which are specified as properties instead (Miao et al. 2005). These properties must dynamically be set to a specific value in order to enact the actual data flow, a task which is performed by the users in runtime through the use of monitor services.

Since the user not the LMS is responsible for managing the data flow, IMS-LD supports a data flow approach which is mainly human-oriented (Palomino-Ramírez et al. 2007), and thus fails in addressing the automation of the data flow, which is the main issue according to the literature: "IMS-LD lacks the management of the data flow" (Peter and Vantroys 2005); "IMS-LD does not support the data flow between tools" (Dalziel 2006); "IMS-LD has no means to specify the relation between artifacts and tools" (Miao et al. 2005); "IMS-LD does not consider the initialization parameters of a service nor whether the output data of a service are going to be exported to any other place" (Wilson 2005); "IMS-LD lacks system-support for addressing the automation of data flow management" (Palomino-Ramírez et al. 2007).

The previous statements show the first evidences of the IMS-LD data flow problem. However, no case studies have been reported in the literature in order to show with real data why the data flow automation is an important issue in LD. For this purpose, next section introduces an authentic case study.

Case Study: A Collaborative Data Flow Scenario

The case study presented in this paper is a shared experience among three Spanish universities which are some of the partners in the major national MOSAIC Learning project (http://mosaic.gast.it.uc3m.es): University Carlos III of Madrid, the Open University of Catalonia and the University of Valladolid. In order that research groups from these universities work together while testing tools from third parties, 14 people participated in this experience in March 2007 playing distinct roles: student, course designer, tutor, and course manager.

The goal of the doctoral students who participated in this experience was the iterative construction of a conceptual map in the topic of grid services and the service oriented computing. The learning design was defined following the approach of Collaborative Learning Flow Pattern (CLFP) (Hernández-Leo et al. 2005) which in this case combines the pyramid, jigsaw and peer-review patterns. The first level of the pyramid comprises 3 phases. In the first individual phase, each student read one out of three selected papers (4 students read the same one) in order to build a first version of the conceptual map. In the second collaborative phase all four students that read the same paper worked together as a group of experts. They made a peer-review of their previous work in order to build a new agreed version of the conceptual map. In the third phase, 3 jigsaw groups of 4 expert students were formed while peer-reviewing remained as the main activity again. In the second level of the pyramid, 2 jigsaw groups worked together in order to improve a new agreed version of the conceptual map. Finally in the third level of the pyramid all the 12 students presented their thoughts.

The overall structure of the aforementioned case study is illustrated in figure 1 as well as in the simplified version of the activity diagram in figure 2a. Note in figure 1 the meaning of the shape and color codes. The triangular, cross, and rectangular shapes corresponds to the members of an expert group, meanwhile the green, red, yellow and blue colors denote the jigsaw group each student belong. Also note in figure 2a that the learning design comprises a learning flow pattern as well as a data flow one.

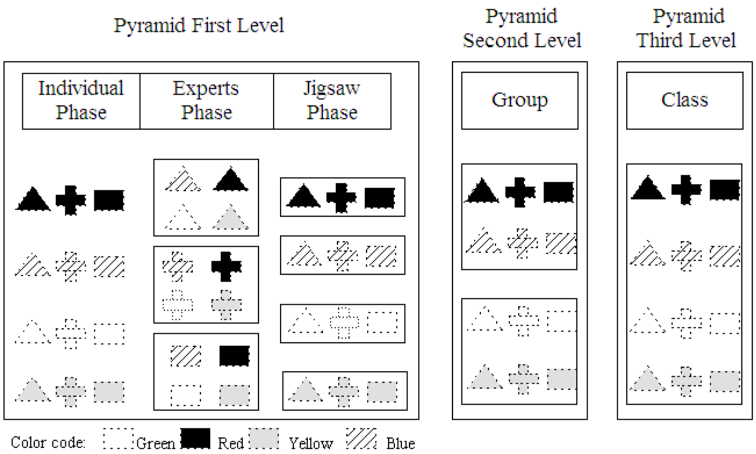
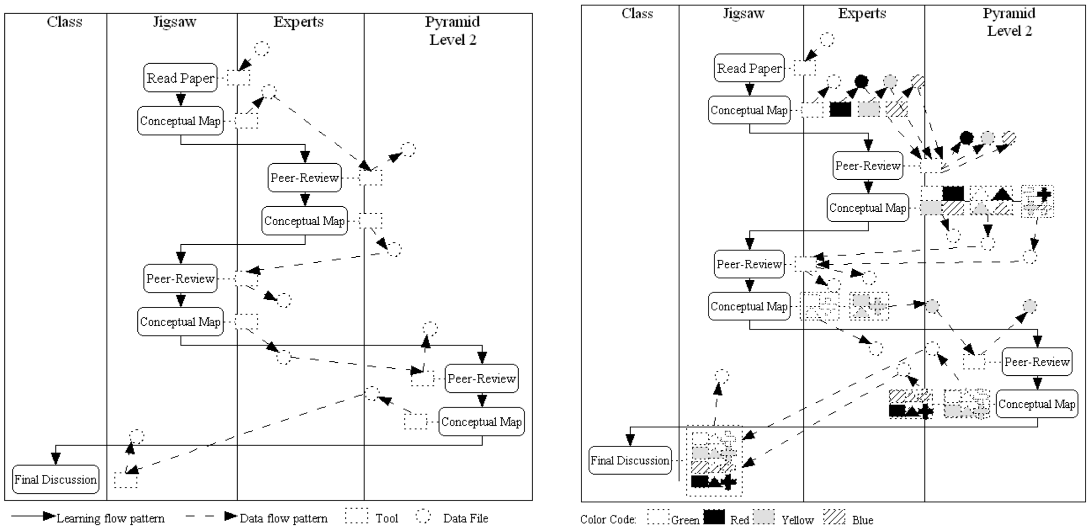


Figure 1: CLFP for the learning scenario.

This case study was not designed in order to analyze the IMS-LD data flow problem. Nevertheless, the aforementioned learning design must be supported by a complex file exchange infrastructure, which is the main issue we deal with in this paper. So, each student must retrieve (individually or in behalf of a group) 25 files and upload 12 files from the shared file system as illustrated by the simplified version of the data flow in figure 2b. Also note that it is specified at the instance-level and not at declarative-level as the data flow pattern in figure 2a. Moreover, it corresponds just for the green-rectangle user while the rest of 12 students that participated in the experience have their own data flow similar to the one shown in figure 2b. Therefore, as we can see, this case study can be considered significative enough within the context of the data flow problem.

Evaluation Methodology

The focus of the case study analysis is based on two data flow perspectives: the users' data management perspective and the course designer's data flow specification one. In order to support our findings, data is gathered from quantitative and qualitative sources, and the analysis follows the principles of the mixed evaluation method proposed in (Martínez-Monés et al. 2003). Table 1 shows a summary of the different data sources considered in the case study which mainly come from questionnaires, e-mail exchange among participants through the experience, as well as from the UoL used in the case study. Two kinds of questionnaires were filled out by the students at the end of the case study: a general purpose questionnaire and a specific one. The former focused on the whole experience while the latter focused on the data flow problem itself. Furthermore, both the course designer and the course managers answered a specific purpose questionnaire at the end of the case study.



(a) **(b)**

Figure 2: A simplified version of the activity diagram of case study comprising the learning flow and the data flow (a) at declarative-level (b) at instance-level (just for the green-rectangle student)

Table 1: Data sources for the evaluation of the case study and labels used in the text to quote them

Data source	Labels
General purpose questionnaire applied to the students after the experience	[student-gquest]
Specific purpose questionnaire applied to the students after the experience	[student-squest]
Specific purpose questionnaire applied to the course managers after the experience	[manager-squest]
Specific purpose questionnaire applied to the course designer after the experience	[designer-squest]
E-mail exchange among participants during the experience (source: student)	[student-email]
E-mail exchange among participants during the experience (source: manager/designer)	[designer-email]
Manifest of the UoL carried out before the experience	[designer-uol]

Results

Table 2 shows the main findings that emerged from the case study, which are described in more detail in this section. Finding 1 supports the relevance of the case study to the IMS-LD data flow problem. As we can see from the students' quantitative data, the managing of files was very important for students and used very frequently through the case study. This finding is also supported with the students' qualitative data, which are also indicative to the problem.

Finding 2 indicates that IMS-LD supports a data flow approach which is error-prone for the users, who are not only concerned in following the instructions for uploading and downloading artifacts to their local file system, but they have to share their artifacts in collaborative activities. The approach used in the case study for this purpose requires that each user be instructed about the right artifact to use from a shared file system. The problem with this approach is that the instructions could be misunderstood, wrongly applied, ignored or forgotten at all (Dillenbourg 2002). Consider the next user comment: "I don't know whether my partner forgot upload his file or uploaded it to any other location, or may be he saved it with other name as it happened to me [student-squest]". This finding is also supported by others qualitative data shown in table 2, as well as by the users' quantitative data: "the 100% of the users state having suffered some kind of error-prone situation during data management [student-squest]".

Finding 3 indicates that the course designer faces also an error-prone situation during data flow specification. Consider the next e-mail exchange between a student and his tutor during the experience: "I'm the yellow-rectangle user, according to the papers summarize I had to read the paper: Grid and Services. However, according to the activity description I'd have to read the paper: Grid Architecture [student-email]". Next, the tutor's answer who is also the course designer: "I'm sorry about the papers assignation. I lost the detail…I've just updated the resources, so the papers assignation already match the files summary with the activity description [designer-email]". It is obviously an error in papers assignation, but interestingly is that the course designer is not only concerned with this issue, but he faces also a data flow consistency problem. On the one hand, the course designer has to specify conditional expressions for each user within the learning design, and on the other hand, he has to instruct each user about the right artifact to use within the activity description. Then, the course designer is not only concerned in the actual data flow, but also in matching the different parts comprising the data flow specification, which is the error-prone situation for the course designer we mean. Note that this finding is also supported by a designer's qualitative data shown in table 2.

Table 2: Main findings of case study including selected supported data and data sources.

Finding	Support	Source
1. The case study is significant and relevant to the problem.	-Students' opinion about the importance of data management [0, 6]: 4.57, deviation 1.99 -Students' opinion about the frequency of usage of the data management [0, 6]: 4.43, deviation: 1.99 -"Files exchange is essential in an experience like this one" -"Data management was used a lot in order to find files from others" -"I had problems to manage paths and file types"	[student-gquest] [student-gquest] [student-gquest] [student-gquest] [student-gquest]
2. The approach employed in data flow management is error-prone for the users.	-Students' opinion about whether the approach employed for data management was error-prone [0, 6]: 3.89, deviation: 1.27 -"I had to put a lot of attention during data management, because it was very easy to be wrong" -"I confused the directory" -"Errors could be greater as more files are shared" -"It is error-prone, particularly at the beginning of the experience"	[student-squest] [student-squest] [student-squest] [student-squest] [manager-squest]
3. The approach employed in data flow specification is error-prone for the course designer.	-"Without enough time for data flow specification, it is in fact an error-prone approach"	[designer-squest]
4. The approach employed in data flow management produces an increase of the cognitive load in the user.	-Students' opinion about whether the approach employed in data management produced an increase of their cognitive load [0, 6]: 3.67, deviation: 1.73 -"It was tedious and messy having to remember each time where and how to save the file"	[student-squest] [student-squest]
5. The approach employed in data flow specification produces an increase of the cognitive load in the course designer.	-"In IMS-LD there is not a common accepted way for data flow management in dynamic groups, so we have to improvise. Obviously, it is an excess of cognitive load" -"There is not a good choice for interaction among groups, so I had to use local personal properties, which increase a lot the complexity" -"The learning flow specification is not too complicated compared to the handle of properties derived from the data flow specification"	[designer-squest] [designer-squest] [designer-squest]
6. The learning design includes instance-level data flow specification.	-24 instance-level properties values corresponding to the green, yellow, blue, red, triangle, rectangle, and cross values are included within the manifest of the UoL.	[designer-uol]

Finding 4 indicates that IMS-LD supports a data flow approach that produces an additional cognitive load for the users. Consider the next user comment: "Data management was not transparent for the student, who had to know and understand how the internal data flow works. It is an important issue, principally for students without technical skills [student-squest]". Again, this finding is also supported by other users'qualitative data shown in table 2, as well as by the next users' quantitative data: "the 100% of users state having suffered some kind of additional cognitive load during data flow management [student-squest]".

Finding 5 indicates that the course designer suffers an additional cognitive load in data flow specification. Consider the next comment from the course designer after asking him about the aforementioned error in papers assignation: "The mismatch in papers assignation, it was just a mistake by myself, due to an excess of cognitive load [designer-email]". Again, this finding is also supported by others designer's qualitative data shown in table 2.

Finding 6 indicates that a collaborative learning design requires an instance-level data flow specification. However, the LD approach requires a separation of the process definition from its execution, so that the UoL could repeat several times for different instance of users and situations (Koper and Miao 2007). Nevertheless, due to the complex interactions between users and artifacts, which are typical in collaborative learning, the declarative-level data flow specification is not enough. Instead, it should be specified at the level of instance, as indicated by the UoL's quantitative data shown in table 2 as well as illustrated in figure 2b.

Previous findings deal with isolated issues of the problem. By combining these findings in a coherent way helps us to understand its relevance. When the collaborative data flow is managed by the users, an additional cognitive load is suffered by them, which is a source of error-prone situations. In a similar way, the concern in matching the different parts of the data flow specification produces an additional cognitive load in the course designer, which is also a source of error-prone situations. Both of them are relevant issues since whether the user get access to the wrong artifact either by a user's or designer's mistake, then the user's learning objectives are not accomplished at all. On the other hand, due to the complex interactions between users and artifacts, which are typical in collaborative learning, the declarative-level data flow pattern is not enough, but an instance-level data flow specification in necessary instead. Nevertheless, merging the generic learning flow with the specific data flow, affects the reusability of the whole UoL, which is a relevant issue too.

Three-Dimensions of the IMS-LD Data Flow Problem

Based on the relation among these findings and literature, three dimensions of the IMS-LD data flow problem have been identified in collaborative learning in order to define the necessary requirements to address it. The first dimension corresponds to the data flow automation problem already reported in the literature. By limiting the user responsibility to upload and download artifacts to his local file system, then the users should not manage the collaborative data flow by themselves, but an automation of the data flow management through the LMS is desired instead, which is our first requirement. The second dimension corresponds to the data flow consistency problem. In order to reduce the designer's cognitive load during data flow specification, it should be specified as a whole not in parts, in the same way than the learning flow one, which is our second requirement. Finally, the third dimension corresponds to the UoL reuse problem. In order to not merge the reusable learning flow specification with the specific data flow one, a separation of the learning flow from the dataflow is our third requirement.

Discussion

There is not a common accepted way in order to specify the collaborative data flow in IMS-LD. In (Koper and Miao 2007) the interaction of artifacts between group members can be specified using global elements while each member of the interaction is modeled in distinct roles. Nevertheless, the designer of the case study opted for using another approach whereby each user is instructed within the activity description about the right artifact to use. Therefore, it could be reasonable to argue that the case study is not valid for a learning design that does not follows the former data flow approach.

Nevertheless it is not the case. In order to specify the share of artifacts among users of distinct groups, when defining a monitor service, the designer has to allow the users to see the properties of other users in a certain role. The problem is that the number of groups is unknown at design time, so groups are actually distinct instances of the same role. Therefore, a monitor service is not enough for the course designer in order to specify the right partner a certain user has to interact. Consider the next designer's comment when asking him about the data flow approach used in the case study: "When using a monitor to see some property, the user has two choices: to see the own properties, or to see all users' properties in the specified role. There is no way to specify that student A only see the file of the student B through the use of global elements…[designer-quest]"

For addressing this issue, the user needs to be instructed within the activity description. But again, this approach requires instance-level data flow specification within the learning design, thus affecting the reusability of the whole UoL. Furthermore, this approach also presents a data flow consistency issue, since the course designer is concerned in matching the different parts comprising the data flow specification. Therefore, we conclude that in collaborative learning, IMS-LD supports distinct approaches for data flow specification but with similar drawbacks. Then, the UoL used in the case study is a valid one, but a comparison between both approaches is out of the scope of this paper.

A Composition-based Solution for the IMS-LD Data Flow Problem

In order to fulfil the aforementioned requirements of the data flow problem while keeping interoperability with IMS-LD, an architectural proposal illustrated in figure 3, which is based on a composition-based approach (Palomino-Ramírez et al. 2007) is next introduced.

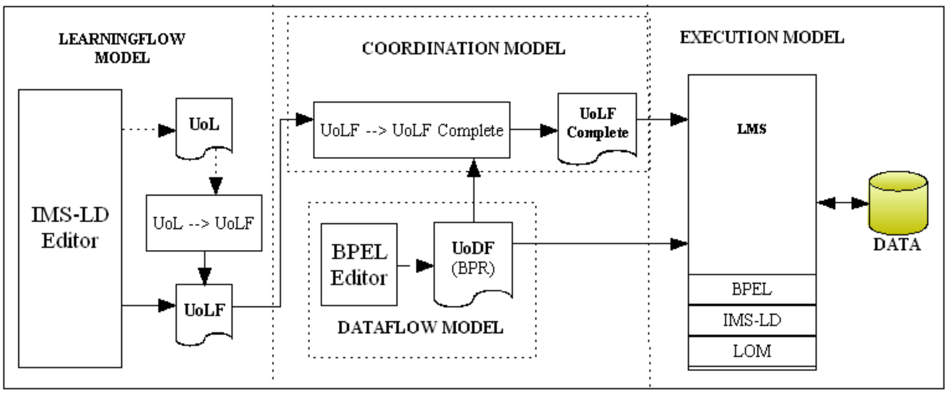


Figure 3: A simplified version of the proposal architecture.

Since IMS-LD fails to specify the data flow in collaborative learning, we propose a separation of the data flow from the learning flow. For this purpose, a standard workflow language such as BPEL is used and a unit of data flow (UoDF), which is just a business process archive (BPR) understandable by a BPEL-compliant engine is created. Then, the learning design is specified in a unit of learning flow (UoLF), which is actually a UoL understandable by an IMS-LD compliant engine that follows the best practices in collaborative data flow specification, which means not specifying the data flow at all. Finally, for coordination of both engines, a coordination model has been developed and a prototype is currently under test and evaluation. Note that it is the BPEL engine not the user, the own responsible for managing the data flow in collaborative activities (the first requirement); the data flow is specified in BPEL as a whole, not in parts (the second requirement); and the data flow is specified separately from the learning flow (the third requirement).

Conclusions and Future Work

In this paper the IMS-LD data flow problem has been analyzed from the perspective of an authentic and relevant case study. Supported with real data, several findings related to the problem in collaborative learning emerged: data flow management and data flow specification are error-prone for both the users and the course designer; both of them suffer an additional cognitive load during data management and data flow specification; and the learning design requires instance-level data flow specification. These findings also helped us to understand the relevance of the problem: a data flow approach which is error-prone for both the users and the course designer may potentially affect the accomplishment of the users' learning objectives; and a learning design which merge the reusable learning flow with the specific data flow affects the reusability of the whole UoL, which are both relevant issues.

Based on the relationship of these findings and literature, three dimensions of the IMS-LD data flow problem have been identified: the data flow automation problem reported in the literature, which is related to the user's collaborative data management issue; the data flow consistency problem, which is related to the issue of keeping consistency in the different parts comprise the data flow specification; and the UoL reuse problem, which is related to the instance-level collaborative data flow specification issue.

ISBN No: 978-1-86220-206-1

Furthermore, these dimensions also helped us to determine the three requirements in order to tackle these issues: it should be the LMS not the user, the responsible for managing the collaborative data flow; the data flow should be specified as a whole not in parts; and finally, the data flow specification should be separated from the learning flow one.

Since IMS-LD fails to specify the data flow in collaborative learning, we propose a separation of the data flow from the learning flow. For this purpose, BPEL is used to specify the data flow within a UoDF while the learning flow is specified within a UoLF. But neither a UoDF nor a UoLF are both new concepts. The former is a BPR understandable by a BPEL-compliant engine while the latter is understandable by an IMS-LD compliant one. Moreover, a UoLF is actually a UoL that follows the best practices in collaborative data flow specification, which means not specify the data flow at all. Finally, for coordination of both engines, a coordination model has been developed and a prototype is currently under test and evaluation. Future work includes evaluation of more case studies in order to validate the proposal solution, as well as to identify limitations and drawbacks. Interestingly is that our proposal which is based on a composition-based approach may be thought as the in-between approach that will provide a framework for future integration of LD and workflow streams.

References

Dalziel, J.R. (2006). Lesson from LAMS for IMS Learning Design. *Proceedings of the 6th International Conference on Advanced Learning Technologies (ICALT 2006)*, Kerkrade, The Netherlands, 1101-1102.

Dillenbourg, P. (2002). Over-scripting CSCL: the risks of blending collaborative learning with instructional design. Three worlds of CSCL. Can we support CSCL, P.A. Kirshner (ed.), Heerlen, Open Universiteit Nederland, 61-91.

Hernández-Leo, D., Asensio-Pérez, J. I. and Dimitriadis, Y. (2005). Computational Representation of Collaborative Learning Flow Patterns using IMS Learning Design. *Journal of Educational Technology & Society*, 8(3), 75-89.

IMS, G. L. C. (2003). IMS Learning Design Information Model. v1.0, Technical Specification.

Koper, R. and Miao, Y. (2007). Using the IMS LD Standard to Describe Learning Designs. Submitted book chapter. Available at http://dspace.ou.nl/bitstream/1820/927/1/UsingIMSLD.pdf. Last retrieved November 6, 2007.

Martínez-Monés, A., Dimitriadis, Y., Rubia-Avi, B., Gómez-Sánchez, E. and de la Fuente-Redondo, P. (2003). Combining qualitative evaluation and social network analysis for the study of classroom social interactions., *Computers and Education* , 41(4), 353-368.

Miao, Y., Hoeksema, K., Hoppe, H. U. and Harrer, A. (2005). CSCL scripts: modelling features and potential use. *Proceedings of the International Conference on Computer Support for Collaborative Learning (CSCL 2005): the next 10 years!*, Taipei, Taiwan, 423-432.

Palomino-Ramírez, L., Martínez-Monés, A., Bote-Lorenzo, M. L., Asensio-Pérez, J. I. and Dimitriadis, Y. A. (2007). Data Flow between Tools: Towards a Composition-Based Solution for Learning Design. *Proceedings of the 7th International Conference on Advanced Learning Technologies (ICALT 2007)*, Niigata, Japan, 354-358.

Peter, Y. and Vantroys, T. (2005). Platform Support for Pedagogical Scenarios. *Journal of Educational Technology & Society*, 8(3), 122-137.

Wilson, S. (2005). Workflow and web services. CETIS White paper. Available at www.e-framework.org/resources/SOAandWorkflow2.pdf. Last retrieved January 22, 2007.

Acknowledgments

This work has been partially funded by the EU Kaleidoscope NoE FP6-2002-IST-507838, Spanish Ministry of Education and Science project TSI2005-08225-C07-04, Autonomous Government of Castilla y León, Spain (projects VA009A05, UV46/04 and UV31/04), Tecnológico de Monterrey Campus Guadalajara, and Fundación Carolina. The authors would also like to thank the rest of EMIC/GSIC research group at the University of Valladolid for their support and ideas and of the rest of MOSAIC Learning project research group for their contributions to this work.