Student activity and profile datasets from an online video-based collaborative learning experience

Estefanía Martín, Manuel Gértrudix, Jaime Urquiza-Fuentes and Pablo A. Haya

Abstract
This paper describes two datasets extracted from a video-based educational experience using a social and collaborative platform. The length of the trial was 3 months. It involved 111 students from two different courses. Twenty-nine came from Computer Engineering (CE) course and 82 from Media and Communication (M&C) course. They were organised into nine interdisciplinary groups. Each group included three or four CE students and eight or nine M&C students. An additional group filmed the making of. This group has only M&C students. Four teachers supervised the trial. The total number of meaningful events was 2984.

Datasets
Two datasets are published
- Activity: contains all the interactions that occur in the learning platform involving both students and teachers.
- Profiles: describe user’s profile including membership, role and background.

Activity dataset
- Location and DOI: ResearchGate http://dx.doi.org/10.13140/RG.2.1.2316.7521 DOI: 10.13140/RG.2.1.2316.7521
- Creator: Universidad Rey Juan Carlos
- Date: October 31, 2014
- Format: CSV file separated by # with header. UTF-8

Profile dataset
- Location and DOI: ResearchGate http://dx.doi.org/10.13140/RG.2.1.2316.7521 DOI: 10.13140/RG.2.1.2316.7521
- Creator: Universidad Rey Juan Carlos

© 2015 The Authors. British Journal of Educational Technology published by John Wiley & Sons Ltd on behalf of BERA. This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.
Introduction
The Juxtalearn Project (2015) looks to promote student curiosity and improve the comprehension of threshold concepts (Meyer & Land, 2003) through the creation of educational videos. The learning process follows a constructionism (Papert & Harel, 1991) approach that can be summarised as “learning by performing.” Students participate in an online reflective learning process (Clough, 2014) that is supported by a social and collaborative platform (Llinás et al., 2014). Students take an active role in their own learning, creating the videos and discussing and evaluating them. Active learning (Bonwell & Eison, 1991) has proven to increase student performance in STEM subjects (Freeman et al., 2014). In any case, teachers have strong control over the way in which the activity is set up, deciding parameters such as group size, duration, content privacy, deadlines, number of expected videos, etc., and to the assessment of the activity such as scaffolding commenting options, rating mechanism, etc.

This dataset offers research opportunities to study collaborative student behaviour in a video-based learning environment. There are several ways to analyse dataset student interactions (Seaton, Bergner, Chuang, Mitros & Pritchard, 2014), including correlation of temporal interaction patterns with deadlines (see Table 1), relationship between levels of participation and ratings (Heo, Lim & Kim, 2010; Vaquero & Cebrian, 2013), difference between inter- and intra-group interactions, how student background affect collaboration patterns (see user profile section) or peer review rating strategies. An example of the outcomes that can be drawn from this dataset is shown in Figure 1. It displays the number of actions (see Table 2) that occurred each day, grouped by voting (fivestar object type) and generic interactions (any other object type).

Furthermore, Haya, Daems, Malzahn, Castellanos and Hoppe (2015) report how they explore the potential of computational methods to support teachers and learners participating in a similar

| Stage                      | Description                                                                 | Who                        | When          |
|----------------------------|-----------------------------------------------------------------------------|----------------------------|---------------|
| Storyline writing          | Short description of the usability concept with the main goal, the approach and a set of references | CE students                | 09/20–10/11   |
| Synopsis writing           | Create a short script including the characters, locations, events and their sequence | M&C students; CE discussion | 10/12–10/25   |
| Script writing             | Developing the narrative, writing the screenplay, and planning the visual and sound production process | M&C students; CE students help | 10/26–11/08   |
| Video preproduction        | Search of multimedia elements that can be included in the video-making step | All students               | 10/25–11/15   |
| Video making               | Making visual and audio production elements                                  | M&C students               | 11/09–11/22   |
| Video editing and postproduction | Video editing and postproduction                                                | M&C students               | 11/23–11/28   |
| Release video final version| Publish the final version of the video in the public side                      | M&C students               | 11/29         |
| Discussion and peer assessment | Students discuss the teaching and production aspects and rate the videos        | All students               | 11/29–12/13   |
collaborative scenario. Although the dataset of this last paper is not publicly available, the learning design and the kind of interaction events (see Table 2) are mainly the same as the one described in this paper.

Research method
During the first semester of year 2013/2014, an educational experience was made at Universidad Rey Juan Carlos (Spain) spanning 3 months. There was an interdisciplinary learning activity between students of Computer Engineering and Media and Communication. The two degrees are taught in different campuses. They were distributed in 10 groups where students from both courses were mixed in each group.

All groups were using an educational social network, called ClipIt, that supports co-creation and peer evaluation scaffolded processes. The last version of ClipIt can be tried out at http://clipit.es. On the one hand, each group has private workspace that includes tools such as a file sharing vault, a messaging service and a discussion forum. This private workgroup space allows students to upload and discuss different elements such as the script, media contents for

---

**Figure 1: Number of actions occurring each day**

![Number of actions occurring each day](image-url)
creating the video and different versions of the educational video. On the other hand, all students involved in the learning experience can comment and rate the videos made by their partners in the public side of ClipIt in order to improve them. Although ClipIt is oriented to video-based learning, any type of other material can be also uploaded and discussed by students.

The main goal of this educational experience was to perform a video that explains usability concepts. Computer Engineering students were “domain experts” in Human Computer Interaction. They explained the usability concept to their partners of the same classroom but from different group. On the other hand, Media and Communication students were in charge of video-producing tasks. The trial was divided into eight stages, as detailed in Table 1. The first column includes the stage name. The second column describes the task assigned in this stage and the expected outcome. The third column reflects who is involved in this stage. Finally, the last column is the stage timespan. Each group produced different materials (storyline, synopsis, script, Table 2: List of object types and actions

| Object type | Action | Event description |
|-------------|--------|-------------------|
| bb_chat     | annotate | A student \( \text{actor} \) sends a private message to another student \( \text{object_id} \) through a chat. |
| blog        | create  | A blog represents a discussion thread. A student \( \text{actor} \) creates, deletes or updates the first blog entry \( \text{object_id} \). |
|             | delete  |                   |
|             | update  |                   |
| file        | create  | A student \( \text{actor} \) adds a new file to the platform. \( \text{object_id} \) indicates the group where the file has been uploaded. |
|             | annotate|                   |
|             | update  |                   |
| fivestar    | update  | A student \( \text{actor} \) rates a video \( \text{object_id} \). Rating field indicates the number of votes. |
| friend      | create  | A student \( \text{actor} \) accepts a friend request from another student \( \text{object_id} \). |
| friendrequest | create | A student \( \text{actor} \) sends a friend request to another student \( \text{object_id} \). |
|             | delete  |                   |
| generic_comment | create | A student \( \text{actor} \) posts a comment on a discussion thread/blog \( \text{object_id} \). |
|             | delete  |                   |
|             | update  |                   |
| group_topic_post | create | A student \( \text{actor} \) posts (or update) a comment on group discussion thread/groupforumtopic \( \text{object_id} \). |
|             | update  |                   |
| groupforumtopic | create | A groupforumtopic represents a group discussion thread. A student \( \text{actor} \) creates, deletes or updates the first forum entry \( \text{object_id} \). |
|             | delete  |                   |
|             | update  |                   |
| images      | create  | A student \( \text{actor} \) adds a new image to the platform. \( \text{object_id} \) indicates the group where the file has been uploaded. |
|             | annotate|                   |
| messages    | create  | A student \( \text{actor} \) sends a message to another student \( \text{object_id} \). |
|             | delete  |                   |
|             | update  |                   |
| pdf         | create  | A student \( \text{actor} \) adds a new PDF to the platform. \( \text{object_id} \) indicates the group where the file has been uploaded. |
|             | annotate|                   |
| ppt         | create  | A student \( \text{actor} \) adds a new PPT to the platform. \( \text{object_id} \) indicates the group where the file has been uploaded. |
|             | annotate|                   |
| user        | profileiconupdate | A student \( \text{actor} \) changes his/her personal information. Ignore \( \text{object_id} \) field. |
|             | profileupdate |                   |
| video       | create  | A student \( \text{actor} \) adds a new video to the platform. \( \text{object_id} \) indicates the group where the file has been uploaded. |
|             | annotate|                   |
| word        | create  | A student \( \text{actor} \) adds a new word file to the platform. \( \text{object_id} \) indicates the group where the file has been uploaded. |
|             | annotate|                   |
multimedia materials and several video versions) for each stage. These materials were uploaded to ClipIt for internal group discussion, except the final video that was shared with all the groups for peer evaluation. The full list of videos is available in the following YouTube playlist: http://goo.gl/aX8HXy.

Activity dataset
The activity dataset includes every meaningful event that occurred during the trial. Each row describes an event excluding the header. Each event has the following fields separated by # character:

<timestamp>#<actor>#<action>#<object_id>#<object_type>#<rating>

- **timestamp**: a UNIX timestamp identifying when the event occurred.
- **actor**: a numeric identifier corresponding to the user that performs an action (see user profile dataset).
- **action**: a verb that described what happened to the object.
- **object_id**: a numeric identifier of the object on which action was taken.
- **object_type**: type of the object. Each type has a different set of actions (see Table 2).
- **rating**: number of votes given to the video on a scale of 1 (poor) to 5 (excellent). This field only applies to a fivestar object type (see Table 2).

Table 2 shows which actions can be performed to each object type. The first column lists object types. The second column includes the possible actions that can be performed in each object type. Finally, the last column describes the event and the meaning of each field.

Profile dataset
The user profile dataset follows a similar format but including different fields separated by # character:

<user_id>#<group_id>#<role>#<background>

- **user_id**: a numeric identifier that identified a user.
- **group_id**: a numeric identifier corresponding to the group he or she belongs to. Teachers were assigned to every group by default.
- **role**: student or teacher.
- **background**: CE (Computer Engineering) or ME (Media & Communication).

Ethical considerations
Datasets have been anonymised in order to ensure that participants cannot be identified. Students uploaded their own videos to Youtube. All participants are of legal age. They knew that they were participants of an educational study and some research studies could be published in several venues. As the videos produced are new educational resources and students are the owners of this material, participants included their own names in the closing credits of the videos. Therefore, it is possible to find the name of participants in the educational material produced, although it is not possible to match the names with the datasets.

Limitations
Lurker behaviour is difficult to characterise as reading events and downloading file is not included in the dataset. Login and logout events were discarded due to inconsistencies in the log file. Student assessment was only based on peer review (see fivestar interactions at Table 2). There was no external grading that measured student’s learning. Finally, students used external mobile messaging and face-to-face communication as well as internal group communication tools. Those interactions were neither recorded nor quantified.
Acknowledgements
The research leading to these results has received funding from the European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement 317964 JUXTALEARN. We would also like to thank Miguel Ángel Gutiérrez Moreno for his support during data wrangling.

References
Bonwell, C. C. & Eison, J. A. (1991). Active learning: creating excitement in the classroom. Washington, DC: School of Education and Human Development, George Washington University.
Clough, G. (2014) Pedagogical framework. Juxtalearn: Deliverable Report D2.3. Retrieved June 15, from http://juxtalearn.eu/images/docs/Juxtalearn_D2.3_WP2.pdf
Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H. et al (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111, 23, 8410–8415.
Haya, P. A., Daems, O., Malzahn, N., Castellanos, J. & Hoppe, H. U. (2015). Analysing content and patterns of interaction for improving the learning design of networked learning environments. British Journal of Educational Technology, 46, 2, 300–316.
Heo, H., Lim, K. Y & Kim, Y. (2010). Exploratory study on the patterns of online interaction and knowledge co-construction in project-based learning. Computers & Education, 55, 3, 1383–1392.
Juxtalearn Project (2015) European Community’s Seventh Framework Programme (FP7/2007–2013) under grant agreement no. 317964. Retrieved June 15, from http://juxtalearn.eu
Llinás, P., Haya, P. A., Gutiérrez, M. A., Martin, E., Castellanos, J., Hernán, I. et al (2014). ClipIt: supporting social reflective learning through student-made educational videos. In C. Rensing, S. De Freitas, T. Ley & P. J. Muñoz-Merino (Eds), 9th European Conference on Technology Enhanced Learning (EC-TEL 2014) (pp. 502–505). Graz, Austria: Springer International Publishing. doi: http://dx.doi.org/10.1007/978-3-319-11200-8_51; 16–19 September 2014 Open Learning and Teaching in Educational Communities, LNCS vol. 8719.
Meyer, J. & Land, R. (2003). Threshold concepts and troublesome knowledge: linkages to ways of thinking and practising within the disciplines. Coventry and Durham, UK: University of Edinburgh.
Papert, S. & Harel, I. (1991). Situating constructionism. Constructionism, 36, 1–11.
Seaton, D. T., Bergner, Y., Chuang, I., Mitros, P & Pritchard, D. E. (2014). Who does what in a massive open online course? Communications of the ACM, 57, 4, 58–65. http://doi.org/10.1145/2500876
Vaquero, L. M. & Cebrian, M. (2013). The rich club phenomenon in the classroom. Scientific Reports, 3, 1174. doi: 10.1038/srep01174.