Learning Analytics with Google Classroom: Exploring the possibilities

Alejandra Martínez-Monés
GSIC-EMIC, Universidad de Valladolid
ETSIINF, Camino de Belén
47011 Valladolid, Spain
+34 983 42300 ext. 5706
amartine@infor.uva.es

Christophe Reffay
ELLIADD, FR-EDUC
University Bourgogne Franche-Comté
57 Av. Montjoux
25042 Besançon, France
+33 0 3 81 65 71 86
christophe.reffay@univ-fcomte.fr

Javier Hoyos Torío
GSIC-EMIC, Universidad de Valladolid
ETIT, Camino de Belén
47011 Valladolid, Spain
+34 983 423969
javierht@gsic.uva.es

Juan A. Muñoz Cristóbal
GSIC-EMIC, Universidad de Valladolid
ETIT, Camino de Belén
47011 Valladolid, Spain
+34 983 423969
juanmunoz@gsic.uva.es

ABSTRACT
Google Classroom (GC) is gaining momentum in the educational milieu, but its functionalities are limited. Learning analytics applications integrated with GC can help to face these limitations, but to reach this aim, developers need access to the data generated by GC's users. This paper reports on the results of an analysis of the existing alternatives to collect data from GC. The study is based on the analysis of the documentation provided by the involved tools. The analysis shows that GC's API is a potential source of data about the activity of the users in GC-enabled settings, but that the information it provides is limited. Further work is needed to explore if Chrome OS synchronization functions can deliver more detailed information about GC usage, thus enabling for more advanced learning analytics applications.

CCS Concepts
·**Applied computing~Education**

Keywords
Learning Analytics; Teachers' adoption; Google Classroom.

1. INTRODUCTION
In spite of the increasing interest in the use of Learning Analytics, there is a scarcity of research on this field for elementary and early secondary education [2]. One feasible reason for this is the overall lack of ICT adoption at these educational levels [11]. An aspect that has shown to be positive for the appropriation of technology by teachers is the use of tools already known by them, such as widespread Web 2.0 tools like the ones offered by Google. The analyses of the educational market also confirm the primacy of Google technologies in the K-12 educational levels [12]. Therefore, the usage of Google Classroom[1] (GC), the platform offered by Google for helping teachers manage assignments in their classes, appears as a good alternative for facilitating the adoption of technology at early school levels.

GC can be complemented with a choice of tools that enhance its functionalities. One of these tools is Hapara[2], a Learning Management System that works together with GC (among others) and offers advanced features for the management of classes, groups, and analytics. According to the documentation available at its website, Hapara is being used at schools, mainly in United States, New Zealand and Australia. Its introduction in Europe is more scarce, mostly based on innovation projects, like E'Pat (Espace Protégé pour l'Apprentissage en ses Traces) [10], where 21 teachers in several primary and secondary schools in France were provided with Chromebooks equipped with Google Classroom integrated with Hapara as a monitoring tool. The main goal of this project was to analyse the consequences of the tools' use in everyday teaching practices and their acceptability by teachers. One of the outcomes of E'Pat was that, in spite of its good properties, Hapara did not fulfil all the needs of the participant stakeholders, especially researchers in LA.

Researchers working in advanced uses of learning analytics combined with GC need to get direct access to the data related to the usage of this platform. The reasons for this need are diverse: a school may be unable to afford the use of commercial tools, but still might be willing to have support for the use of GC; researchers may want to have a register of the actions performed by the participants to carry out advanced analysis, or to deliver

[1] Google Classroom web site: https://edu.google.com/products/productivity-tools/classroom/ Last visit 1-9-2017

Conference'10, Month 1–2, 2010, City, State, Country.

DOI: http://dx.doi.org/10.1145/12345.67890

[2] Hapara web site: https://hapara.com/ Last visit 1-9-2017

new kinds of solutions, etc. However, up to our knowledge, there has not been any learning analytics approach based on direct usage of data provided by GC.

The fact that Hapara's "Teacher dashboard" is able to provide detailed information of the usage of the documents by the students drives us to ask whether this information can be obtained directly from GC. To answer this question is the main goal of this paper. From the perspective of a researcher in learning analytics, there are potential benefits in knowing which features are provided by GC without the need of using third-party tools. Direct access to data from GC would allow to adapt learning analytics proposals to the particular needs of GC's users. Taking into account the aforementioned numbers of uptake of GC at early school levels, this could eventually help face the current lack of adoption of LA by practitioners, especially in K-12 education.

The structure of the document is as follows: we start outlining the main features provided by GC and Hapara. Then, Section 4 analyses the different alternatives to collect data from GC. Section 5 presents three scenarios to illustrate the benefits and limitations of the analysed alternatives. Then, we discuss the results of the analysis and finally, we present the conclusions and propose the future work lines.

2. GOOGLE CLASSROOM

Google Classroom (GC) is a collaborative tool for teachers and students that builds on existing Google technologies, mainly Google Docs, Google Drive, and Gmail. A teacher working with GC can setup classes to which students can join using a code provided by the platform. Within a course, teachers can: create and distribute assignments and attach different types of documents to them; monitor the state of the students' submissions; and provide feedback to the students once the assignment is turned in. Teachers can grade the assignments and manage the grades using Google Spreadsheets.

GC is not a learning management system (LMS) [7]. It lacks many of the features that LMSs provide to their owners, especially those related to the automatic management of the lists of students and gradebooks. This may explain the fact that institutional uptake of GC was slow until 2016 [1].

GC is especially suited for its use with Chromebooks, -notebooks equipped with the Chrome OS as its operating system-. A survey to more than 1000 K-2 teachers and administrators in USA showed a high increase in the adoption of GC with Chromebooks, with 42% of the teachers naming Chromebooks as the devices provided by the schools, due to their easy of use and affordability, and G-suite and Google Classroom being selected by 78% of the respondents as their favourite tools in the classroom [6].

GC strength may be related to the capability to manage Google's ecosystem in ways that fit well with teachers and students' existing practices [7]. According to a recent study [8], the main benefits of GC are the management of assignments and the possibility for teachers of keeping the work organized by class. The main drawbacks pointed out by users are that GC is by design focused on assignments, and therefore, it does not provide many of the features provided by actual LMSs, such as quizzes and forums. It also has limitations regarding the ways a teacher can keep track of the students' progress in the assignments. For example, GC requires a teacher to go into each student folder (if they are working on individual documents) in order to be aware of their performance. These limitations among others explain the need of third-party tools that extend the functionalities of GC to provide schools and teachers a better user experience. As aforementioned, Hapara is one of such tools. Next section describes Hapara, focusing on the features that extend GC.

3. HAPARA AND THE TEACHER DASHBOARD

Hapara is an educational platform that integrates Google Apps and improves how teachers can manage the classroom. It organizes the student and the teacher in the Google environment—docs, comments, blogs, calendar, portfolios-, to reflect school and class structure.

Hapara consists of three tools that give teachers more insight of students' work and activity.

1. **Hapara Workspace** consolidates all the various apps, resources and activities into a series of cards in a particular workspace. Teachers can add learning goals and rubrics, and they can customize particular students' workspaces. They can also organize collaborative workspaces around ability-level, language barriers, learning or conceptual goals, and more. Activities contain a built-in workflow with embedded due dates and submission instructions. Teachers can see a summary of their entire class's activity and track the assignment submission status and individual activity levels. The workspace can also be used with teachers in a professional learning setting.
2. **Hapara Highlights** enables teachers to track students' web activity and makes it visible and immediately actionable for teachers and administrators.
3. **Hapara Dashboard** enables teachers to track student activity across all sites within the school's cloud-based platform. Teachers can organize these individual activity dashboards by class, by the specific application and by a specific student.

All these functions rely on a back-end architecture where Hapara communicates directly with GC to retrieve data about students' activity into the system. However, up to our knowledge, there is no document describing the architecture of Hapara and its relationship with GC. This makes it difficult to understand the means by which these two applications communicate and complement each other, and therefore, leaves the question of how much data can be retrieved by a third-party about students' usage of GC open. Section 4 section

delves into this issue.

4. APROACHES TO RETRIEVE DATA FROM GOOGLE CLASSROOM

We explore in this different approaches that could be used to retrieve data from GC for research purposes. This analysis is based on the documentation provided by the systems analysed, and should therefore be considered as an initial attempt to understand the problem and its possible solutions.

We start reviewing the solutions that could be provided directly by GC, and after that, we review other possibilities that involve the instrumentation of the systems to get the data directly from the users' interactions with the computers on which GC is installed.

4.1 Alternatives based on Google Classroom functionalities

First of all, GC does not provide a log or database access that can be retrieved to be analysed by third party tools. Therefore, alternative approaches to the use of logs have to be proposed to use GC together with Learning Analytics solutions.

GC provides basic monitoring support. If GC is installed as part of Google Suite for education, the site administrators have access to the following data[3]: Number of active users per week; Number of active classes during the last 14 days; number of courses created; number of publications created by teachers and by students. For a concrete user, the number of classes and publications created, as well as the time of the last activity performed in GC. However, this information is provided to administrators as front-end, in ways that can not be processed by third-party tools. This impedes, for example, the possibility of giving personalized feedback to teachers and students or to adapt to the specific needs that may found in different contexts.

GC offers also monitoring support for teachers: once teachers have shared a document, they can see which users are working on it in a particular class. This feature opens a new tab for each student. Therefore, it does not provide an overall view of the classroom activity in a single screen, making it difficult for teachers to identify students that need special attention, or to derive general conclusions about the classroom's work.

A third alternative is to use GC's API[4]. This API can be used to integrate third-party applications, such as those based on learning analytics, in GC. GC's API defines several entity types that enable this integration, like the ones devoted to the creation and management of courses (Course, Alias, Invitation), the management of participants (Student, Teacher, User Profile), the description of the assignments created by teachers (CourseWork), and the students' submissions (StudentSubmission). Table 1 shows the attributes of the latter entity. Some of these attributes allow to keep track of the moment when the students submit and update their assignments (creationTime, updateTime) as well as the state of theses assignments (state). Table 2 describes the possible states defined for an assignment, which match the types of actions allowed in GC, where students can create, turn in and reclaim an assignment, and teachers can return an assignment to a student. The attribute SubmissionHistory[] yields the history of the states and grades a submission has gone through. In summary, the entities defined by the API allow to keep track of part of the activity that the students and teachers carry out in GC and could be used by an external tool to provide some degree of personalized feedback. However, more detailed information, such as the actual actions performed by the students on the objects related to an assignment, is not provided. This level of information may be necessary for some kinds of learning analytics applications that need to get access to this information to provide personalized feedback. Therefore, we can conclude that GC's API provides limited detail about the users' activity within a course, that can be enough for providing some degree of extra functionality but that may not suffice for getting more advanced feedback.

Table 1. Main attributes of the entity StudentSubmission

Field	Description
courseId	Identifer of the course
courseworkId	Identifier of the course work this corresponds to
id	Classroom-assigned identifier for the student submission. Unique for the course
userId	Identifier of the student that owns this submission
creationTime	Creation time of the student's submission (unset if the student has not submitted anything)
updateTime	Last update time of this submission (unset if the student has not accessed the submission).
state	State of the submission (See Table 2)
late	Boolean filed indicating if the submission is late
draftGrade	Optional field with a provisional grade only visible and modifiable to teachers.
assignedGrade	Optional field with a grade which is only modifiable by teachers.
alternateLink	Absolute link to the submission
courseWorkType	Type of course work (an assignment, a short answer question or a multiple choice question).
associatedWithDeveloper	A boolean value that controls access rights to this submission.
SubmissionHistory[]	History of the submission including the state and grades of the assignments.
Content	The content of the submission. It is an Union field, which type depends on the value of courseWorkType.

Table 2. List of possible states of a Submission

States	Description
NEW	The student has never accessed this submission
CREATED	The submission has been created
TURNED_IN	The submission has been turned in to the teacher
RETURNED	The submission has been returned to the student (for further review)
RECLAIMED_BY_STUDENT	Student chose to "unsubmit" the assignment (they may want to improve it, or for any other reason)

[3] https://support.google.com/edu/classroom/answer/7283376?hl=en Last access 2-9-2017

[4] https://developers.google.com/classroom/ Last access 2-09-2017

4.2 Alternatives based on external instrumentation

Other approaches to get information about GC usage are based on the instrumentation of the system, so that the activity performed by the users is captured by third-party tools. One possibility could be to use Google Analytics together with Google Classroom, but according to its documentation, Google does not provide this possibility by default[5]. In order to use Google Analytics, we could try to instrument the particular GC website, as described in [3], but even if this is done, the information provided by Google Analytics may be not well-fitted for many learning analytics purposes. Google Analytics is a business oriented tool that provides information about network traffic in a Website: total number of visits to a page, their duration, number of pages visited by a user, time spent by a user in each page, session length, how do users interact in the pages, etc. It is possible to get real-time data (users that are viewing a page in a given moment, etc.). Even if not provided by Google Analytics by default, it is also possible to identify the user that access the website and thus, provide information based on each user, like shown in [3]. However, the information given by Google Analytics is based on low-level events, such as accesses to the website. It is difficult to connect this information to actions related to the course activity (e.g., students' submissions, readings, etc.). This is a major obstacle for the use of Google Analytics as a source of data, as learning analytics tools require the understanding of upper-level actions so as to produce meaningful feedback for their users.

Finally, we could consider potential alternatives that do not depend on Google Classroom, but on the operating system where it is running. For example, Chromebooks (or any dispositive running Chrome OS) allows to sync devices among them, which means that they keep track of the state of the browser. If an external tool could retrieve this synchronization information, and could match it with contextual information about the course and its characteristics, it could give a good idea of the activities carried out by a teacher or a student in the system. According to the documentation about the synchronization of files[6] , it could be possible to identify what the users are doing with the files (and eventually, to get access to the owner of the files, keep track of their advances in the system). Due to technical limitations of this research, this possibility has not been further explored in this report, but remains as part of the future work derived from it.

5. ILLUSTRATIVE SCENARIOS

In this section, we describe hypothetical scenarios to illustrate how the different solutions discussed in the previous section could be integrated as part of a system using learning analytics. The scenarios that are presented below include to a range of the possible stakeholders and goals that can be identified for learning analytics [4], and aim at showing the benefits and limitations of the different alternatives.

Scenario 1: Google's front end to help an administrator: Sally is a school administrator. She works for a Secondary School that has decided to acquire Chromebooks and open a Google Classroom institutional account. In order to inform the School Board, Sally is interested in knowing the level of usage of the system by the school overall and by the different teachers, classified by department and by course. With the information provided by Google's administrators front-end and some post-processing, she can provide these quantitative measures to the members of the School Board. Sally also would like to know whether the usage of the system is perceived as positive or negative by the teachers and the students, and even if there is any observable effect on the students' grades after the introduction of the platform. The system does not provide this information, but she does not worry too much about that, because she is in continuous contact with the staff and will be able to get their opinions thanks to their formal and informal meetings. Regarding the grades, she thinks she will be able to carry out a rather rough analysis thanks to the statistics facilitated by the official system used by all the schools in the district to keep students' grades. Anyhow, she is aware that there are so many factors that can influence the grades that it would be very difficult to justify a change in these grades just performing a basic study like the one she is able to do with her resources.

Scenario 2: GC's API to help a teacher manage groups in a flipped classroom: Javier is an innovative teacher of Science, known by his colleagues for being an early adopter of technologies but also a very self-reflective professional. He is worried by the fact that his students do not carry out the assignments he has carefully designed to help them understand his subject. He promotes group work and carries out a type of flipped classroom where he asks the students to do some individual homework and share it in the classroom in groups. He is using GC, as he finds this system very well suited to his purposes, but he losses a lot of time organizing the working groups at the beginning of the class. Javier would like to know right before each class which students have (and have not) submitted their individual part, so as to have this information into account when configuring the groups at the beginning of the class. Thanks to his contacts in the University, he has collaborated in a project where a developer has built a simple tool based on GC's API, which takes into account the information about the students' submissions provided by the API and configures the groups automatically, attending to some parameters introduced previously by Javier. This is very important for Javier, who can concentrate in attending his students' questions instead of losing his time organizing groups. He would also like to know the percentage of contribution of the students to the group work, but this is not possible to get with the GC's API. He has set up another project with his colleagues at the University to find out whether the sync functionalities of the Chromebooks used in the classroom can help to that aim.

Scenario 3: Google Analytics for a researcher: Pedro is a researcher on learning technologies who is interested in the time-based behavior of students when they work in assignments. He would like to use Javier's classroom as a case study, as his teaching method obliges the students to submit many assignments of different types, both individually and with different group sizes. Pedro thinks that instrumenting the GC website where Javier is running his course with Google Analytics will give him information of the time-based distribution of accesses to the site by the different users. However, he will need to ask Javier about the course design and assignments to be able to match the observed behavior with the different kinds of tasks performed by the students, and he will only be able to analyze the behavior of the students while they work on the GC site (not on other platforms, as they normally do). In order to carry out this study, Pedro has asked for permission to the School Principal and to the

[5] See https://productforums.google.com/forum/#!msg/google-education/ObmNOjuB3Yg/VmbBA8ZDucwJ

[6] https://developer.chrome.com/apps/syncFileSystem#event-onFileStatusChanged

parents of the students (as they are minors), making sure that he will anonymize all the data collected, and no information about any individual student behavior will be shared with anybody out of the classroom. Pedro thinks that having a platform that helped him match the students' actions with their meaning and facilitating anonymization would be of a great help for his research.

6. DISCUSSION

The two previous sections have presented initial insights into the question of whether the usage of GC can be monitored in ways that can be later processed by third-party tools to provide personalized and adapted learning analytics solutions. As shown in the review, GC does not provide logs, and it cannot be used with Google Analytics, and the administrator's monitoring support, although useful for a general overview, does not allow to personalize learning analytics to the needs of different educational contexts. More concretely, it does not provide actionable visualizations of the system, i.e., data easy to interpret by the participants and that invites them to reflect and intervene on the learning scenarios.

The review shows that one approach to achieve actionable visualizations could be based on the information provided by the GC's API. The types of questions that could be answered by this hypothetical learning analytics system would be restricted to the kind of learning activity promoted by GC, which is focused on students' submission of assignments. However, even within this scope, it could be interesting to integrate this information in a system that could be aware of the teachers' lesson plans such as the design-aware learning analytics proposed by [9]. The envisioned system could take into account teachers pedagogical intentions and compare them to the state of the interaction in a given moment, like illustrated in the second scenario presented in the previous section. Taking into account the possibilities of the GC's API, this output is restricted to questions such as whether students have submitted an expected assignment, or if they have received teachers' feedback on time to carry out the next activity in the plan, etc. These analyses would not be fine-grained, as GC's API does not give very detailed information about the actions of the participants in the documents. However, it could be still interesting to see how the more coarse-grained information available can be merged with other sources of data (maybe the content of the submissions or, in blended environments, information given by the participants) to reach useful solutions.

Finally, in cases where more detailed information is needed regarding the actions of the students, the synchronization facilities of Chromebooks could be the answer. However, this solution is restricted to schools that implement GC with this technological setup.

7. CONCLUSIONS AND FURTHER WORK

The use of Google Classroom is becoming widespread, and this could help foster teachers' uptake of technology, specially at K-12 levels, where the challenge is bigger. However, GC provides a limited functionality, and the existing third-party applications integrated with it do not provide the kind of actionable, context-aware information that could be very valuable to complement the functionalities provided by GC. Learning analytics applications can help to fill this gap, but for doing that, it is necessary to collect data from GC at a meaningful level of abstraction.

We provide in this paper a review of different alternatives that exist to get data from GC's usage. The main conclusion obtained is that the most promising alternatives are, firstly, to exploit the possibilities of the API provided by GC; and, secondly, to analyse whether it is possible to use the synchronization facilities provided by Chrome OS to instrument the system and get information about the activities of the participants in the system.

This study has important limitations, as it is exclusively based on the analysis of the documentation provided by the analysed tools. An obvious line of future work is to carry out case studies where the alternatives analysed are implemented and showcased, in order to gain further insight into their benefits and limitations. A second line of future work is to design learning analytics scenarios that make use of the data-collection alternatives discussed in this paper, and analyse whether the information that could be offered by these alternatives can improve the services provided by existing tools, such as the aforementioned Hapara.

A final word remains to point out that any attempt to get data about users' usage of the system, especially in school contexts, needs to be framed in a very solid ethical framework, where all the data privacy and ethical issues are considered and taken into account before starting the project. The fact that GC builds on Google applications adds complexity, as the relationship between what it is offered as front-end and how it is managed as data in the back-end of the system are not clear, and could make it not compatible with rules at the educational systems [5].

8. ACKNOWLEDGMENTS

This research has been partially supported by the Junta de Castilla y León (VA082U16), and Ministerio de Economía y Competitividad (TIN2014-53199-C3-2-R, and TIN2015-71669-REDT). This work has been partly supported by a grant as invited professor for two months offered by the University of Franche-Comté to the first author and by the Ministry of Education of Spain. The authors also want to thank the reviewers that contributed to improve the final version of this paper.

9. REFERENCES

[1] EdTech 2016. Apple, Microsoft, Amazon and Google Are Fighting a War for the Classroom. *edutechnica. EdTech Talk and Analysis*.

[2] Ferguson, R., Brasher, A., Clow, D., Cooper, A., Hillaire, G., Mittelmeier, J., Rienties, B., Ullmann, T. and Vuorikari, R. 2016. Research Evidence on the Use of Learning Analytics: Implications for Education Policy. (2016).

[3] Filva, D.A., Guerrero, M.J.C. and Forment, M.A. 2014. Google Analytics for time behavior measurement in Moodle. *Information Systems and Technologies (CISTI), 2014 9th Iberian Conference on* (2014), 1–6.

[4] Greller, W. and Drachsler, H. 2012. Translating learning into numbers: A generic framework for learning analytics. *Journal of Educational Technology & Society*. 15, 3 (2012), 42.

[5] Lindh, M. and Nolin, J. 2016. Information we collect: Surveillance and privacy in the implementation of Google Apps for Education. *European Educational Research Journal*. 15, 6 (2016), 644–663.

[6] Morrison, N. 2017. Google leapfrogs rivals to be classroom king. *Forbes/Education/#onCampus*.

[7] Mullaney, T. 2015. Google Classroom is not an LMS. It's better. *Sustainable teaching.*

[8] Pappas, C. 2015. Google Classroom Review: Pros And Cons Of Using Google Classroom In eLearning.

[9] Rodríguez-Triana, M.J., Martínez-Monés, A., Asensio-Pérez, J.I. and Dimitriadis, Y. 2015. Scripting and monitoring meet each other: Aligning learning analytics and learning design to support teachers in orchestrating CSCL situations. *British Journal of Educational Technology*. 46, 2 (2015), 330–343.

[10] Schneewele, M. and Reffay, C. 2017. *Bilan du projet EPA'T : Espace Protégé pour l'Apprentissage en ses Traces*. Université de Franche-Comté, Université Bourgogne Franche-Comté

[11] Somekh, B. 2008. Factors affecting teachers' pedagogical adoption of ICT. *International handbook of information technology in primary and secondary education*. (2008), 449–460.

[12] Walters, N. 2017. Google Takes Classroom Lead, While Apple and Microsoft Fight for Second Place. *The street*.