# Use of a CSCW platform by three different groups: trace analysis according to Activity Theory.

Jean Simon, François-Marie Blondel

# Use of a CSCW platform by three different categories of people: trace analysis according to Activity Theory

**Jean SIMON, Equipe GRRAPELI, IUFM de la Réunion, FRANCE**
**jean.simon@reunion.iufm.fr**

**François-Marie BLONDEL, ENS de Cachan, INRP, FRANCE**
**francois-marie.blondel@inrp.fr**

*Abstract : In this paper, we analyze the traces of the activity on a CSCW platform. For that we gather the traces in the higher level shared folders (hlsf ). Based on Activity Theory, the hlsf permit to distinguish between the groups and the goals they pursue.*
*The first group that we will study is constituted of preservice teachers who use the platform to pool and share teaching resources. The second group is constituted of students who use the platform to prepare an examination and the third one is constituted of researchers who use the platform to work together.*
*We will show that according to the groups and their goals the observed activities are not the same.*

Keywords : trace analysis, CSCW, Activity theory, teachers training

## 1. Introduction

We study here the activity of three different categories of users on two CSCW platforms BSCW (Bentley  & al, 1997).
The first category consists of preservice teachers (PE2s for "professeurs des écoles 2è année") who learn how to teach, they have passed successfully the contest to become a primary school teacher. The second consists of students (PE1s for "professeurs des écoles 1è année") who prepare this contest. The platform used in their case is a BSCW platform managed by the training teachers school of  La Réunion  (IUFM for Institut universitaire de formation des maîtres). It is important to note that we study their work when they are alone on the platform and that there is no trainer associated with their work. We have showed in (Simon, 2009a) and (Simon, 2009b) that it makes a difference in terms of results.
The latter category consists of a community of students and researchers. The platform used then is a BSCW platform managed by the ENS (Ecole Normale Supérieure) of Cachan. This is the reason why we will call this category "ENS".
These three categories had a lifespan very different when we have registered the traces of their activity (july 2009). The two first had access to the platform only during one academic year (2008-2009) when they were in formation in the IUFM, while the students and the researchers have access since September 2004 to the platform of ENS.

The reasons to use a platform of groupware are different from one category to another. For PE2s, the platform serves to pool and share resources to make class. For the PE1s, the objective is rather to share everything that can help them to success in the contest. At the ENS of Cachan, the students work on the platform at the demand of their trainers because most of them are not always present on campus. For researchers, the platform serves to preserve and exchange on key documents and events and manage a large number of collective actions, including projects that involve other remote teams in France or abroad.
It is noticed that if the students and the researchers of the ENS and the PE2s of the IUFM can find that it is very beneficial for them to work together, it is not the case of the PE1s since they will be in competition at the time of the contest.

To study this activity, we analyze the traces which these users have left on the platform by using the Activity Theory (Engeström, 1987) because these three types of activity are characterized as well as by the groups which are at their origin as by the goals they pursue.

In the next section, we quickly present the adopted methodology and, in section three, the obtained results. In the conclusion, we reconsider these results.

## 2. Methodology

The principles and the reasons of the methodology that we employ here can be found in (Simon et al, 2008) and (Simon 2009a). Briefly, we study the traces that were left on platforms by combining them in units that make sense.

For each category, we have the totality of the traces which its users have left on the platform. These traces are like photography, taken at a given time, of the activities taking place on it.

At the time of the analysis, there were 166893 of them left by the researchers and the students on the platform of the ENS of Cachan and 775750 on the platform of the IUFM of La Réunion. But for the latter, the traces do not concern only the categories of users PE1 and PE2 but also secondary school teachers, trainers,...

In most researches, trace analysis consists simply in counting the traces. The problem is that, by doing this way, a lot of information is lost. These traces reflect various activities and this is the reason why it is necessary to be able to gather them according to these activities. In particular, we must be able to define who has left them to pursue what purpose. We find, in the background of this approach, the Activity Theory.

In order to be able to recover these two parameters, the group and the goal, we, thus, gather the traces in *hlsf* (higher level shared folder) (Simon, 2009,a). A *hlsf* is a folder created and shared by a group of users in order to achieve their goals. A *hlsf*, thus, includes the traces corresponding to different objects (sub-folders, documents, URLs, ...), various actions on these objects (creation, reading, modification,...) and the group of members associated with it.

The use of the hslf as unit of analysis makes it possible to study the traces left on the platform according to these groups and of their objectives. In particular, this has allowed us to study the activity of PE2s during three years, distinguishing whether they were alone (Simon, 2009a) or with a trainer (Simon, 2009b).

In what follows, we analyze the *hlsfs* shared by each of the three categories: PE2s, PE1s and researchers.

## 3. Results

### 3.1. *Number of hlsfs*

Table 1 number of *hlsfs*

| Number of *hlsfs* | PE2 2008-2009 | PE1 2008-2009 | ENS 2004-2009 |
|---|---|---|---|
| number of *hlsfs* | 72 | 22 | 102 |
| number of users | 175 | 165 | 325 |
| number of *hlsfs* for one user | 0,41 | 0,13 | 0,31 |

Table 1 indicates the number of *hlsfs* per category of users, the number of users and the number of *hlsfs* brought back to the number of users.

The PE2s created the biggest number of *hlsfs* but we will see later that these *hlsfs* were not the most active, moreover, compared to the previous years (Simon, 2009, a), this number is going down.

The PE1s have the lowest number of *hlsfs* both in absolute terms (22) or reported to the number of PE1 (0.13). This is understandable when we know, as we have noted, that the PE1s will be in competition. This will also appear in most of the next tables.

Regarding the ENS category, it can be wondered why they do not have created more *hlsfs* because their activity take place over 5 years whereas the activity of the PE took place only over one year. The answer holds, perhaps, in the fact that all the *hlsfs* are associated with an activity and that a larger number of years does not imply necessarily a larger number of activities, a same activity being able to proceed over several years.

### 3.2. *User involvement in several hlsfs*

Table 2 : participation to one or more *hlsf* .

| participation to one or more *hlsf* . | nb of PE2 2008-2009 | % | nb of PE1 2008-2009 | % | ENS 2004-2009 | % |
|---|---|---|---|---|---|---|
| 0 | 12 | 6,86% | 39 | 23,64% | 3 | 0,92% |
| >=1 et <10 | 121 | 69,14% | 126 | 76,36% | 311 | 95,69% |
| >=10 | 42 | 24,00% | 0 | 0,00% | 11 | 3,38% |
| TOTAL | 175 | 100,00% | 165 | 100,00% | 325 | 100,00% |

Table 2 indicates the number of *hlsfs* in which the users take part. It is noted, in all categories, that a majority of them take part from one to ten *hlsfs*.

We note what has been reported previously: It is among the PE1 that we find the greatest number of non-participants. Nearly 1 PE1 out of 4 does not belong to any *hlsf*.
In contrast, only 1% of the people of ENS are not involved in any *hlsf* and they are also a minority to participate in more than 10.

It is among the PE2 that the greatest number of users taking part in more than 10 *hlsfs* is found. That raises the question to know if all their *hlsfs* are justified or if so some of them could not be gathered.

### 3.3. *Roles*
Table 3 : Roles

| Roles | Number of PE2 | % according to the nb of PE2 members | Number of PE1 | % according to the nb of PE1 members | Number of people of ENS | % according to the nb of ENS members |
|---|---|---|---|---|---|---|
| leader | 41 | 25% | 17 | 13 % | 22 | 7% |
| moderator | 49 | 30% | 17 | 13% | 58 | 18% |
| producer | 62 | 38% | 18 | 14% | 95 | 29% |
| reader | 133 | 82% | 106 | 82% | 316 | 98% |
| inactive | 30 | 18% | 23 | 18% | 6 | 2% |

| member of one hlsf at least | 163 | 100% | 129 | 100% | 322 | 100% |
|---|---|---|---|---|---|---|
| Total number | 175 | 107% | 165 | 128% | 325 | 101% |

Table 3 indicates how the users have distributed the roles among them. We distinguish five different roles. The leaders who have created at least one *hlsf*, the moderators who have created at least one subfolder in these *hlsfs*, the producers who have deposited at least one document in these *hlsfs*, the readers who have read at least one document in these *hlsfs* and, finally, the inactives who, as the name suggests it, did not take part in the life of the *hlsfs*.
For each category, the percentages are reported to the number of members of the category who take part in at least one *hlsf* and not to the total number of members of the category.

We can see that according to the categories to which they belong, the users have distributed the roles differently.

### 3.3.1. *Organization (leader and moderator)*

We see that there is no real leadership among the PE2s, one PE2 out of four has created a folder, whereas there's only one out of eight for the PE1s and only one out of sixteen for the researchers. This concept of leadership is to be taken with caution for the reason that it is not because someone launches a *hlsf* that he is necessarily the leader of this *hlsf*. This is the case, for instance, at the ENS where one researcher has created 52 *hlsfs* out of 102 and most of them have been created at the demand of his colleagues.

The tendency remains the same for the moderators but with some nuances. It is for the PE2s that we find the greatest percentage of moderators. For the PE1s , the number of moderators is  equal to the number of leaders. Furthermore, a closer analysis shows that this is the same persons who have created the *hlsfs* who have been brought to create sub-folders.

At the ENS, the number of moderators is much bigger than the number of leaders. This suggests a greater division of labour in this category. Another explication is that, once the general framework is decided through the creation of the *hlsf,* more people create subfolder according to the life of the projects.

### 3.3.2. *Production*

It is noted that the number of producers reported to the number of member is weakest among the PE1s. That can be explained, again, by the fact that they are in competition.

This percentage is nearly 38% among the PE2s and 29% among students and researchers of ENS. For the latter, this number can appear relatively weak and we would have thought that over the 5 years all members would have, more or less, taken part in the production of documents. This could be explained by the fact that all the students are not always constrained to produce, in the same way some researchers produce little, often because these means of communication and exchange do not belong to their professional practices.

We will reconsider the production in the last point of this section because the number of producers by *hlsf* is what characterizes the best these *hlsfs*.

### 3.3.3. *Use/reading*

Concerning the use, the readings, we note that almost all the students and researchers were interested in what there was on the platform (98% have read at least one document) but only 8 PE out of 10 (PE1 or PE2)

### 3.3.4. *Partial conclusion*

Concerning PE2s, the figures seem to indicate that there is no really leadership or people charged with the organization. We can also wonder whether there were, a priori, organization of those *hlsfs*.
The same phenomenon of strong imbalance between production and use, already reported in (Simon, 2007) and (Simon 2009a) is confirmed here. If we assume that PE2s must share documents to ease their tasks while they are in charge of a class, we find that 44% of them do not really respect the rule because they take without giving (82% -38%).

However, it is at among the PE1 that this phenomenon of "lurker" is the strongest: only 14% of them deposit documents whereas they are 82% to read. The fact that they are 64% to take without giving seems to be consistent with the fact that they ate in competition.

Finally, for students and researchers of the ENS, there are more moderators than leaders. This suggests a greater division of labour than in the previous categories: the management of collective projects implies a distribution more constrained of the professional frameworks.
For them too, there is a gap of 68% between the number of producers and the number of readers. But given the objectives given by them to the platform, this gap is less troublesome because the activity does not aim to ease the tasks of each other but rather to broadcast the information.

## 3.4. *Activity*

We have seen in previous tables, the number of participants and the roles they were given. We want to go further and discover up to what point this participation is important.

Table 4 : Activity

| activity of the *hlsf* . | PE2 2008-2009 | PE1 2008-2009 | ENS 2004-2009 |
|---|---|---|---|
| Number *of hlsf s* | 72 | 22 | 102 |
| Number of PEs | 175 | 165 | 325 |
| Average number of members for one *hlsf* | 18 | 15 | 12 |
| Average number of documents for one *hlsf* | 4 | 6 | 44 |
| Average number of documents produced by one member | 0,24 | 0,44 | 3,55 |
| Average number of readings for one *hlsf* | 18 | 50 | 137 |
| Average number of readings for one *member* | 1 | 3 | 11 |

In table 4, we observe the activity within *hlsfs* shared by each category. To measure this activity, we evaluated the average number of members in the different types of *hlsfs*, the average number of documents produced and the average number of readings done.

Thus, we see that if the average number of members per *hlsf* is not very different from one community to another, the productivity of its members is significantly higher among students and researchers (44 documents produced by *hlsfs*) than PE2 (4 documents produced by *hlsfs*) and that this remains true even if we reduce the work of researchers to one academic year (44 / 5 = 8.8).

If we consider, moreover, the average productivity per member, it reveals that students and researchers of ENS are the most productive. However, as noted in Table 3, all the members are not producers and we analyze more precisely this in the next point.

Concerning the readings, by contrast, we see that, if we take into account the lifespan of the *hlsfs*, the PE1 read the most, 3 readings by each member, while PE2 are reading, on average, only 1 document and the students and researchers 2 (11/ 5)

Furthermore, we also find that all members do not read all the documents and, this, whatever the community. Thus, there could have been 72 readings by the PE2s (18x4) and there were only 18, 549 readings by the students and the researchers while there has been only 137. It is among the PE1 that the number of real readings, 50, is closest to the number of possible readings, 90 (6x15).

### 3.5. *Production*

The number of producers in a *hlsf* is one of the criteria which make it possible to define the type of activity within this *hlsf* (Gerard, 2009). In one hand, more there are producers, more there are exchanges and there could be cooperation or collaboration within those *hlsfs*. On the other hand, we can assume that *hlsfs* where we find only one producer are rather intended for dissemination of information.

When we consider the *hlsfs* according to the number of producers, we obtain the following table.

Table 5 : number of  *hlsf* according to the number of producers who work in it

| number of  *hlsf* according to the number of producers who work in it | PE2 2008-2009 | | PE1 2008-2009 | | ENS 2004-2009 | |
|---|---|---|---|---|---|---|
| | nb hslf | % | nb hslf | % | nb hslf | % |
| 0 producer | 7 | 9,72% | 2 | 9,09% | 5 | 4,90% |
| 1 producer | 27 | 37,50% | 14 | 63,64% | 62 | 60,78% |
| 2 producers and more | 38 | 52,78% | 6 | 27,27% | 35 | 34,31% |
| total | 72 | 100,00% | 22 | 100,00% | 102 | 100,00% |

In table 5, we note that there is the biggest number of producers in the *hlsfs* shared by the PE2. Almost 53% of their *hlsfs* have 2 or more producers. This number confirms those of previous years (Simon, 2009a). This reflects the fact that the PE2 want to pool and share resources to make class. In addition, as we have seen, it is in the *hlsfs* shared by PE2 that there is the lowest number of documents, we can deduce that the PE2s are the least productive.

In *hlsfs* shared by PE1s, as in *hlsfs* shared by students and researchers, there is generally one single producer. We can thus assume that we are less in exchange of information between members than in dissemination of information from one member to the others.

However, the reason is very different from one category to another. Whereas, in the case of the researchers, it is the diffusion of information, to state to the others what each one does, which is at the origin of the creation of the *hlsfs*, in the case of PE1s, the small number of

producers is explained by the fact that they are competing with each other as we have already seen in previous points.

To conclude this point, we can say that :
- the PE2s are, at most, in cooperation rather than collaboration in the sense of (Dillenbourg & al, 1996)
- for the two other categories, the platform is primarily used to disseminate information but not for the same reasons.

### 3.6. *Organisation*

| organization of the *hlsf* . | PE2 2008-2009 | PE1 2008-2009 | ENS 2004-2009 |
|---|---|---|---|
| Number *of hlsf s* | 72 | 22 | 102 |
| A: Average number of documents for one *hlsf* | 4,46 | 6,55 | 44,17 |
| B :Average number of subfolders for one *hlsf* | 0,44 | 1,18 | 10,11 |
| Percentage of *hlsf s* with one level at least of subfolder | 13,90% | 40,91% | 56,86% |
| Average nb of documents for one *folder (A/(B+1))*[1] | 3,10 | 3,00 | 3,98 |

It is noted that, naturally, more there are documents more there are subfolders, but it is also noted that we find the phenomenon of overorganization announced in several of our previous articles. We call "overorganization" the fact that there are very few documents on average per folder. We showed that in 2006-2007, the *hlsfs* shared by PE2s when they are between peers had an average of 3 documents per folder, that remains true over the next three years and also for the *hlsfs* that PE2s have shared with trainers (Simon, 2009b).
We note, here, that the community of the PE1s and the community of students and researchers do not derogate from this rule because we obtain an average of almost 3 documents per folder for the PE1s and an average close to 4 documents per folder for the researchers.
We put quotes at "overorganization" because we can wonder whether this phenomenon does not come rather from a lack of organization of the data rather than an overorganization as reports (Reffay & al, 2008)

## 4. Conclusion

Even if we did not refer to it explicitly, all our work has for background the Activity theory and the different heights of the triangles of Engeström (Engeström, 1987).
We find, thus, the concepts of subject, of goal (pool and share resources to make class, to prepare the contest, use the platform to diffuse its work), of community (PE2, PE1, students and researchers), of tool (BSCW) but also of division of labour and in a less way of rules.
To be able to make this analysis, we could not simply count the traces left on the platform. Those traces should have been gathered in *hlsfs*. The *hlsfs* make it possible to analyze more precisely what occurs on a platform and to distinguish between the different activities which are held there and the groups which are at the source of these activities. This enabled us to

---

[1] As the *hlsf* is also a folder we integrate it in the average (B subfolder +1 hlsf).

show that according to the groups and the goals which they pursue the activities are not the same.

For PE2, the objective is to pool and share resources to make class. This explains that there are more producers than in other *hlsfs*. Nevertheless, we find that the activity in their case seems to be less intensive, and more diffuse than in other communities:
- more *hlsfs*
- more organizers(leaders, moderators)
- less documents and readings per *hlsf*.

This phenomenon is still reinforced by the fact that BSCW is not the only tool they use for pool documents. In an investigation carried out in July 2009 (Gérard, 2010), they indicate that they also use the email and USB keys to share resources.

There is also a great difference between the number of producers and the number of readers. In the investigation cited above, almost all PE2s wish to continue to pool after training. We can wonder about the possibility of pooling whereas 44% of people do not do it.

For the PE1s, the use of a CSCW platform may seem contradictory in the sense that the contest puts them in competition with each other. This can explain :
- why they are few to take part in a *hlsf*,
- why they are also very few to produce,
- why the readers read so much.

A possible strategy, when we want to pass a contest, may be to take as much information as possible and to give as little as possible. The results found here are going in this way.

That, by contrast, raises the question of the motivations of the producers. Indeed, they seem to work against their interests because they deliver information to people who do not give them back. This unequal exchange should hurt them a priori but, in fact, this is not the case. On the 165 PE1 who have done the contest only 25 have succeeded (15%), while on the 18 PE1 producers, shown in Table 3, who have done the contest, 6 were successful (33%).

Concerning this last point, it is however necessary to avoid making a cause and effect link. It is not because that these 6 PE1 were producers on the platform that they had more chance to succeed at the contest but more probably it is because these 6 PE1 were more motivated that they produced on the platform AND succeeded at the contest.

For the category students and researchers of the ENS, the tool, BSCW, seems to be less a tool for collaborative work or even cooperative work in the strict sense of (Dillenbourg et al, 1996) as a tool for disseminating information to colleagues. We can not really speak of exchange when there is one producer only per *hlsf*.

Another figure that is going in this sense is the very low number of notes. A note on BSCW allows, among others, to comment a document. They are fewer than 300, all types of notes confused, in the space shared by them.

Concerning the division of labour, the work of the researchers is more organized than in other categories: most producers than moderators and more moderators than leaders.

But for this category it is clear that the analyse can go deeper. The next one will have to split this category in two subcategories, the hlsfs shared by students and trainers and the hlsfs shared by the researchers for projects.

The analysis of the traces is thus optimized by the fact that we gather them in *hlsfs*. It can still be improved by a multimodal analysis: social networks analysis of the groups associated with the *hlsfs*, titles analysis of the names of the folders,… (Simon, 2009b). But nevertheless, if

we want that the vision is more complete, it is still necessary to supplement it by investigations of the users. It is what we did in July 2009 (Gerard, 2010).

## Bibliographie

Bentley R., Appelt W., Busbach U., Hinrichs E., Kerr D., Sikkel K., Trevor J., Woetzel G. (1997). Basic Support for Cooperative Work on the World Wide Web. *International Journal of Human Computer Studies: Special issue on Novel Applications of the WWW, Spring 1997,* Academic Press*,* Cambridge, vol. 46, n°6, p. 827-846.

Dillenbourg, P., Baker, M., Blaye, A. & O'Malley, C.(1996) he evolution of research on collaborative learning. *In E. Spada & P. Reiman (Eds) Learning in Humans and Machine: Towards an interdisciplinary learning science*. (Pp. 189-211). Oxford: Elsevier.

Engeström Y. (1987). *Learning by expanding: An Activity-Theoretical Approach to Developmental Research*. Orienta-Konsultit Oy.

Gerard JP, (2009)   Analyse des réseaux sociaux associées aux dossiers partagés par des professeurs des écoles stagiaires, *TICEMED*,  Milan, 2009

Gerard JP,(2010) TCAO : Représentation des usages d'une plate-forme par les professeurs des écoles stagiaires, *soumis à JOCAIR*, Amiens, 2010

Reffay C. Basque C (2009) Echanger pour apprendre en ligne In *STICEF* Volume 15, 2008, pp 41-58

Simon J,  (2007) Teaching resource pooling and sharing by the primary school teachers trainees , *ICOOL* 2007, Penang, Malaysia, 2007

Simon J,  Gerard JP,, Thevenin C., (2008), Dossiers partagés par les stagiaires avec ou sans formateur à l'IUFM de La Réunion : Analyses des traces. In STICEF Volume 15, 2008,

Simon J, (2009 a) Three years of teaching resource sharing by primary school teachers trainees on a CSCW platform, CSCL 09, *Computer Supported Collaborative Learning Practices - CSCL2009 Conference Proceedings*, *ISLS* , pp 267-271, Rhodes, 2009

Simon J, (2009 b) Three years of use of a CSCW platform by the preservice teachers and the trainers of the Reunion Island teacher training school, ICALT 09, *Proceedings of the 2009 Ninth IEEE International Conference on Advanced Learning Technologies - Volume 00,*  pp 637-641, Riga, 2009