Journal of Universal Computer Science, vol. 18, no. 19 (2012), 2602-2624
submitted: 30/3/12, accepted: 6/9/12, appeared: 1/11/12 © *J.UCS*

Obtaining Requirements for Designing a Tool to Support Distributed Development

José Luis Hernández
(University of Castilla-La Mancha, Ciudad Real, Spain
JoseLuis.Hernandez@uclm.es)

Aurora Vizcaíno
(University of Castilla-La Mancha, Ciudad Real, Spain
Aurora.Vizcaino@uclm.es)

Ismael Caballero
(University of Castilla-La Mancha, Ciudad Real, Spain
Ismael.Caballero@uclm.es)

Gabriela Aranda
(National University of Comahue, Neuquén, Argentina
garanda@uncoma.edu.ar)

Abstract: The Distributed Software Development involves various challenges, many of which are related to the lack of trust experienced by team members since they often do not know each other personally. Other problems which are already considered to be classic are related to communication, coordination and collaboration. Bearing in mind that providing information about co-workers may increase both the team spirit and the team members' confidence in each other, we have carried out surveys to discover what information might be useful in attaining this objective, and to reduce those problems related to the three 'c's. The results of the questionnaires have been analyzed from different points of view, first by differentiating the respondents' roles and then by analyzing their specific use in communication, coordination and control activities. The results of this analysis have allowed us to obtain the requirements needed to design a tool with which to support global software development. The implementation of this tool is also described in this paper.

Keywords: Trust, Distributed Software Development (DSD), Global Software Development (GSD), Tools for GSD
Categories: D.2.2, D.4.4

1 Introduction

Recent decades have witnessed the emergence of a strong trend in the development of software through virtual teams, in which their members are located in different cities (known as Distributed Software Development, DSD) or even countries (then known as Global Software Development, GSD). These approaches imply various advantages, such as a reduction in production costs [Conchúir et al. 2009], principally when some of the companies involved are located in countries with low salaries, taxes etc. However, this dislocation also has negative consequences. One of the principal

problems arises as a result of the lack of face-to-face communication, which is the richest type of communication in comparison to alternative settings [Carmel 1999; Oshri et al. 2007]. This makes understanding among co-workers more difficult, and consequently impoverishes the social relationships. GSD team members therefore often complain about a loss of team spirit and a lack of trust [Conchuir 2010], and this is an important problem since collaboration in virtual project teams relies largely on interpersonal trust [Rusman et al. 2010].

Moreover, the costs of coordination and control often increase in global settings [Espinosa 2004]. Once again, the lack of trust influences these aspects, as is stated by [Corbitt et al. 2004; Jarvenpaa et al. 2004] who explain that interpersonal trust between team members is broadly acknowledged to benefit coordination. Furthermore, cultural diversity creates new challenges since the team members often have different native languages, cultures, beliefs and customs. It is therefore advisable to make certain knowledge about the co-workers' backgrounds available to all of them in order to facilitate communication and understanding [Shachaf 2008].

People in virtual teams form an impression of their co-workers based on the information that they may obtain [Walther 2005]. Various studies have been carried out in order to discover the influence of different information types, such as text, video or audio, on the richness of trust formation [Bos et al. 2002]. However, it would appear that it is not still clear whether specific information "does the trick" in professional settings [Rusman et al. 2010]

Nevertheless, this information can be provided by using different means such as story-telling, role playing games, team-building exercises or personal profiles. In professional contexts, the last means (personal profiles) is that which is most broadly used, and contains static or dynamic information about a person's identity [Danis 2000]. The use and goal of the information available in these profiles depends on the context in which they are to be used.

Considering that it is rather important to share and know information about co-workers in order to improve the trust level and, consequently, communication, coordination and control, we believed that it would be of great interest to discover what information software engineers would like to obtain about other members. Therefore, a survey was conducted in order to discover what information people working in a distributed setting usually know, and what information they would like to know but do not have. The conclusions obtained from this survey have helped us to design the profiles of a tool that we have developed to support GSD in an attempt to reduce certain communication, coordination and control problems.

The reminder of this paper is structured as follows: Section 2 describes how the survey was designed and how it was distributed to the different companies. Section 3 provides a description of the results obtained from the analysis of the data. We would like to clarify that some of these results were previously presented in [Aranda et al. 2010]. The main differences between the aforementioned work and those presented here are that the number of surveys formerly completed was 23, in comparison to the 36 to be found in this paper. Moreover, the focus used in the ICGSE paper for analyzing the data was totally different to that used in this paper, as is explained in Section 3. Then, in Section 4 the tool designed after considering the requirements obtained from the survey and the revision of literature is described. Finally, in Section 5, our conclusions and future work are outlined.

2 Designing and Conducting the Survey

In order to discover what information software engineers who work in a distributed setting should know about their co-workers, or at least that which they do not know but would like to know, a survey containing questions about these items was design. The survey was sent to several companies, who were asked to respond to it. The initial idea was that the respondents would be DSD project members playing different roles within the projects.

The structure of the survey was as follows: First, some questions about the respondents' experience in distributed or global software development projects were included, such as:

- Distributed Software Development (DSD):
  - Please indicate the number of projects on which you have worked in a distributed manner.
  - Please indicate the number of months that you have worked on distributed projects.
  - In which cities were the other sites located?
- Global Software Development (GSD):
  - Please indicate the number of projects on which you have worked in global software.
  - Please indicate the number of months that you have worked on global software projects.
  - In which countries were the other members located

The respondents were also asked about the language used to communicate, along with the role that they often play when they are working on a distributed or global project.

A table containing various data items was then presented, and the respondents had to answer the questions about each item [see Table 1].

	I usually know this:		I think this information is:					It is useful/not useful in:		
Data Item	YES	NO	VU	U	N	NV	NU	Cm	Coor	Ctrl
Other person's first name	X		X					X	X	
Other person's surname	X		X					X	X	
Other person's nickname		X				X				
Gender	X						X			

Table 1: Table model of the survey

There are two kinds of questions: 1) those that require a yes or no answer, such as: do you usually know this data item with regard to the co-workers on the other sites?; and 2) those questions in which people have to indicate how important the data item is for them. In this case, the possible answers are shown in Table 2.

Very Useful	VU	4
Useful	U	3
Normal	N	2
Not very useful	NV	1
Not at all useful	NU	0

Table 2: Possible answers to determine the usefulness of data items

The items included in the questionnaire are listed in [Appendix 1]. They were obtained after interviewing five (5) engineers with experience in global software development. The list was completed with other data found in personal social networking profiles such as facebook or linkedin.

An example of the table, as it appears in the survey, was shown in Table 1, along with a possible answer from a stakeholder.

The respondents were also requested to indicate to what extent each data item could be considered as useful for the communication (Cm), coordination (Coor) or control (Ctrl) tasks. They were therefore additionally requested to classify each data item into one of the three categories (communication, coordination and control) with regard to the use made of them in their company. For example, the first row of Table 1 signifies that the respondent considers "*Other person's first name*" to be a data item that is usually known by him/her. Moreover, s/he thinks that this data item is very useful in general, and particularly for communication and coordination.

At the end of the survey there were some questions regarding the tools that software engineers often use for communication and project management, along with others regarding the information related to the project that they often share.

The survey was sent to seven companies working in distributed and global software development. Although the engineers were heartily encouraged to complete the survey, only 36 questionnaires were received [see Table 3]. The names of these companies have been omitted for reasons of privacy.

Company	Total Received Surveys	Nationality of respondents
A	7	Spanish
B	9	Spanish
C	4	Brazilian
D	5	Spanish
E	7	Argentine
F	2	Mexican
G	2	Indian

Table 3: Number of surveys from each company

The questionnaires were transcribed and prepared for analysis from different perspectives so that the respondents could be grouped by business and according to their roles in the projects.

3 Analyzing data from the surveys

In an earlier study of the questionnaires, the data items were classified into different categories according to the percentage of respondents who knew each piece of data and how useful that data was considered to be at a work level [Aranda et al. 2010].

This paper intends to focus on another different aspect. Now, the studies have been carried out from three points of view. Our objective is to analyse the data set of user knowledge and the information lacks that they have in the area of DSD and GSD.

We shall first study the influence that the users' roles have on information needs [Sect. 3.1]. That is to say, we shall analyse which data each role considers to be useful. Once this data is known, it is going to be analysed whether this information is available to them. We shall then analyse those data items which are considered to be useful in communication, coordination and control [Sect. 3.2]. This study will be carried out on all the respondents with the objective of discovering those data items which are considered to be useful for each of the categories. Finally, the tools used for communication between the stakeholders were analysed [Sect. 3.3]. The intention of this analysis is to discover whether any direct relationship exists between the use of communication tools and the quantity of data items which is known.

The following values will be calculated for each data item:

- General Usefulness.
- Known data.
- Rate to which the data item is Useful for Communication, Coordination and Control.

The first step was calculating the level of General Usefulness of each piece of data in the professional working environment. This value signifies how useful each data item is considered to be. It is obtained by adding the weights of the responses concerning the usefulness of each data item (ranged between 0 and 4, see Table 2) and dividing this by the maximum possible value in order to obtain a percentage. For example, for the data item "Other person's first name" the sum of the responses to all the questionnaires is 137. The maximum that could be obtained is 36*4 = 144. The value of general usefulness for this data item is, therefore 137/144= 0.9514 (95.14%).

The data items are classified according to their percentage of usefulness in the following ranges [see Table 4]:

General Usefulness				
Not al all useful	Not very useful	Normal	Useful	Very Useful
0%	0% - 25%	25% - 50%	50% - 75%	75% - 100%

Table 4: General Usefulness

The percentage of users who know each data item (Known Data) was also calculated, which allowed us to discover less well-known data items.

Finally we calculated the percentages of usefulness for communication, coordination and control, obtained from the columns Cm, Coor and Ctrl of the questionnaires. These values provided us with data items which are considered to be useful for different categories.

Each of the analyses is explained in detail as follows.

3.1 Data Analysis by role

Before carrying out the data analysis by role, the questionnaires were grouped by company and a study was carried out of the percentage of data items known by each respondent. Therefore, Table 5 shows the percentage of data items known by each respondent in the seven companies (A-G). For example, in the case of Company A where there are 7 respondents. The first value (46.51%) in Table 5 shows that this person knows less than half of the data items, whilst the last questionnaire in this company indicates that the respondent knows 60.47% of the data items, i.e., more than half. The ratio which appears in the last row is the average percentage of data items known in the company.

		Companies						
		A	**B**	**C**	**D**	**E**	**F**	**G**
Surveys	**1**	46.51%	37.21%	30.23%	32.56%	37.21%	76.74%	62.79%
	2	32.56%	25.58%	25.58%	55.81%	76.74%	58.14%	23.26%
	3	55.81%	65.12%	32.56%	44.19%	55.81%		
	4	46.51%	58.14%	51.16%	37.21%	72.09%		
	5	48.84%	41.86%		34.88%	79.07%		
	6	48.84%	81.40%			27.91%		
	7	60.47%	81.40%			58.14%		
	8		53.49%					
	9		44.19%					
	Ratio	46.94%	47.79%	32.67%	39.43%	50.96%	66.16%	33.94%

Table 5: Ratio of known data items

In this study we realized that within the same company there will be some users who know more data items than others. This signifies that they have more information at their disposal than other colleagues. We wished to discover whether this difference was influenced by each respondent's role, and the following step was, therefore, to classify respondents' answers according to their roles in the company. The 33% of the respondents were project managers, whilst the rest were analysts and programmers. According to the data 92.31% of the project managers knew a percentage of data items which was superior to the average amount of data items known by all the respondents. The result of this analysis is fairly clear, and demonstrates that the project managers know a higher amount of data items. Because of this, we decided to carry out a more in-depth study of the data according to each role.

The specific study of the roles was carried out on all the questionnaires, without considering to which company they belonged. We sought to discover what information needs the users had with regard to their roles.

The respondents were grouped in 3 groups, according to their roles. Given that some users played various roles, we took the highest role played by each one, which resulted in:

- *Programmers* (P): respondents who are programmers and are neither analysts nor project managers.
- *Analysts* (A): respondents who are analysts and are not project managers.
- *Project Managers* (M): respondents who are project managers.

Figure 1 shows the distribution according to these roles.

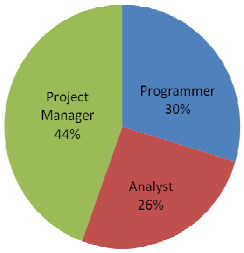

Figure 1: Distribution of participating roles in the study

The analysis was therefore carried out on these three study groups (Programmers, Analysts and Project Managers).

The research questions for this role analysis were:

- Which data items do programmers consider useful but do not know?
- Which data items do analysts consider useful but do not know?
- Which data items do project managers consider useful but do not know?

The following aspects were considered:

- A data item is considered to be useful if the percentage of usefulness is above 50%, i.e., the study group considered it to be useful or very useful.
- A data item is not sufficiently well-known if the percentage of "Known Data" is less than or equal to 50%, thus signifying that it is known by less than half the users.

In order to answer the first question, *which data items do programmers consider useful but do not know?* An analysis on the group of programmers was carried out. We therefore studied the data items that the programmers considered to be useful, but which a high percentage of them did not know.

The most relevant data items in this study are shown in Table 6. For example, the programmers considered the data item 41 "*Technologies in which s/he has experience*" to have a level of usefulness of 78.13% (very useful), yet only 50% of them knew this information. Comparing this with the information provided by the project managers, it was discovered that they do know these data. For example, the data item referring to *Languages spoken* is considered to be very useful (71.88%), and only 25% of programmers actually know this information, whilst the percentage of project managers who know it is 75%.

id	Data Item	**Known by** Programmer	**Considered Usefulness by** Programmer
41	Technologies in which s/he has experience	50.00%	78.13%
42	Programming languages in which s/he is an expert	50.00%	78.13%
15	Languages spoken	25.00%	71.88%
35	Planned meetings (date, time and place)	25.00%	68.75%
16	Level of English	12.50%	65.63%

Table 6: Data items considered as useful by Programmers, although not widely known.

In order to answer the second question, *which data items do analysts consider useful but do not know?*, a similar analysis to the previous one was carried out, but in this case the study group was made up of the analysts. The most significant data items are those shown in Table 7. For example, the analysts consider data item 42, "*Programming languages in which s/he is an expert*" to be fairly useful (57.14%), but only 28.57% of them know this information.

id	Data Item	**Known by**Analysts	**Considered Useful by** Analysts
42	Programming languages in which s/he is an expert	28.57%	57.14%
14	Other person's stress levels	14.29%	50.00%

Table 7: Items of Data considered as useful by Analysts, although not widely known.

Finally, in order to answer the third question, *which data items do project managers consider useful but do not know?* We carried out a study on the project managers. It was discovered data items that project managers wished to know, since they considered them to be useful, but were only known by a low percentage [see Table 8].

id	Data Item	**Known by** Manager	**Considered Useful by** Manager
36	Other person's availability (available, having ...	33.33%	60.42%
14	Other person's stress levels	33.33%	64.58%
13	Other person's emotional state	25.00%	56.25%
23	Preference in carrying out determinate role	41.67%	68.75%

Table 8: Items of Data considered as useful by Project Managers, although not widely known.

Of these data items we could mention data item 36, "*Other person´s availability*", which project managers considered to be useful (60.42%) but which was only known

by 41.67% of them. It is fairly logical that project managers should wish to know different data items to those of the other two roles. For example, data item 23 "*Preference in carrying out determinate role*" was not considered to be excessively relevant by programmers and analysts, whilst project managers did consider it to be so.

After carrying out the role study, gaps appeared. That is to say, there were data items that certain users considered to be useful with regard to their role, but only a low percentage knew them. We therefore considered that these data items should be included in the profiles of the tool that we intended to develop, and we obtained the following requirements:

The first requirement that should be considered is that the tool should capture and show programmers the following information:

- Technologies in which s/he has experience.
- Programming languages in which s/he is an expert.
- Languages spoken.
- Planned meetings (date, time and place).
- Level of English.

The second requirement would be that the analysts would have access to the following data:

- Programming languages in which s/he is an expert.
- Other person's stress levels.

The third requirement would be to show the project managers the following data:

- Other person's availability.
- Other person's stress levels.
- Other person's emotional state.
- Preference in carrying out determinate role.

This first analysis has, therefore, produced these three requirements which will be dealt with later in Section 4.

3.2 Identification of data items considered to be useful for communication, coordination and control

Given that communication, coordination and control are the most crucial activities in DSD, the following phase was to analyze those data items which are considered to be useful for the various activities but which are, nevertheless, known by few users. To do this, we considered the following questions:

1) Which data items are considered to be useful for communication and are unknown by the majority of users?
2) Which data items are considered to be useful for coordination and are unknown by the majority of users?
3) Which data items are considered to be useful for control and are unknown by the majority of users?

The study was carried out on a single group made up of all the respondents. How many users considered each piece of data to be useful for communication, coordination and control was calculated. Of those data items which were considered to be most useful for the various categories, we concentrated on those which were

widely unknown. After that, the differences between General Usefulness for work and usefulness for communication, coordination and control were compared.

In order to answer the first question: *Which data items are considered to be useful for communication and are unknown by the majority of users?* The data that the respondents considered to be useful for communication were analysed. Of this data items set we concentrated on those data items which a high percentage of users do not know. For example, data item 3, "*Other person's nickname*", which can be seen in Table 9, is not considered to be of a high level of usefulness as regards that which is Generally Useful for work, and is known by a low percentage of respondents (38.89%). However, it is considered to be of a high level of usefulness for communication. This signifies that despite not being considered to be directly useful at a work level, it is important for communication, and it would therefore be appropriate to consider it in order to favour communication between stakeholders. Consequently, it can be considered as a requirement for a tool whose purpose is to favour communication, along with the other data items shown in Table 9.

ID	Data Item	General Usefulness	General Useful Level	Known	Useful for Communication
3	Other person's nickname	43.75%	Normal	38.89%	77.14%
12	Other person's personality	49.31%	Normal	44.44%	77.14%
13	Other person's emotional state	40.28%	Normal	13.89%	77.14%
14	Other person's stress levels	51.39%	Useful	27.78%	74.29%
9	Cultural aspects of other person's country	54.17%	Useful	50.00%	71.43%
36	Other person's availability (available, having breakfast, in meeting, etc.)	65.28%	Useful	47.22%	71.43%
5	Age	39.58%	Normal	30.56%	65.71%
27	Personal e-mail address	36.11%	Normal	25.00%	60.00%
8	Birthday	24.31%	Not very useful	13.89%	51.43%
20	Companies in which s/he has worked	41.67%	Normal	25.00%	51.43%

Table 9: Useful data items for communication tasks

In order to answer the second question: *Which data items are considered to be useful for coordination and are unknown by the users?* An analysis similar to that shown above was performed, but by analyzing the data items that the respondents considered as useful for coordination. A list of those data items which were known by a low percentage of the respondents was obtained [see Table 10]. For example, data item 14, "*Other person's stress levels*" is considered to be useful for coordination (65.71%) and is known by very few respondents (27.78%).

According to the respondents, knowledge of the data items shown in Table 10 might improve project coordination.

ID	Data Item	General Usefulness	General Usefulness Level	Known	Useful for Coordination
14	Other person's stress levels	51.39%	Useful	27.78%	65.71%
36	Other person's availability	65.28%	Useful	47.22%	62.86%
21	Years of work experience	50.00%	Normal	38.89%	57.14%
12	Other person's personality	49.31%	Normal	44.44%	54.29%
23	Preference in carrying out determinate role	45.14%	Normal	22.22%	54.29%

Table 10: Useful data item for coordination

Finally, we attempted to answer the third question: *Which data items are considered to be useful for control but are unknown by the majority of respondents?* It is important to mention that there were a lower number of respondents for the control analysis. This was perhaps owing to the fact that control functions are not as evident as those of communication and coordination, and is carried out by more specific roles such as project managers. We therefore carried out the control study on the project managers.

The data items that the project managers considered to be useful for control but were insufficiently known are listed in Table 11. Those which stand out are data items such as personality or emotional state, which are the most difficult data items to obtain.

ID	Data Item	Useful for Work	Usefulness Level	Known	Useful for Control (Project Manager)
12	Other person's personality	49.31%	Normal	44.44%	83.33%
13	Other person's emotional state	40.28%	Normal	13.89%	66.67%
21	Years of work experience	50.00%	Normal	38.89%	58.33%
14	Other person's stress levels	51.39%	Useful	27.78%	50.00%

Table 11: Useful control data items for Project Managers

This study led to the emergence of three new requirements, in addition to those mentioned in the previous section. The fourth requirement is, therefore, that the tool should facilitate the information shown in Table 9 in order to favour communication. The fifth one is that the tool must provide the data items concerning coordination, shown in Table 10. And the sixth requirement is to that the tool should supply to project managers with the information shown in Table 11.

3.3 Analysis of tools used

The principal objective of this last analysis was to obtain the requirements with which to design a tool to support DSD. We therefore wished to analyse which communication tools were used by the respondents. Our objective here was to discover whether the use of a particular tool had influenced the results obtained in the survey. Therefore the following two questions were considered:

1) Which communication tools are most frequently used by the companies?

2) Do any significant differences exist in the use of communication tools which might justify the difference in data items knowledge that exists between the companies?

We analyzed the data items from the questionnaires in which the respondents stated which communication tools they used. In the questionnaires they were asked about the use of telephones, Messenger, e-mail, OCS, and a determinate field dominated as 'others' in which the respondent could state which tools were used, amongst which we found skype, video-conferences and Google talk.

In order to answer the question: *Which communication tools are most frequently used by the companies?* An initial general study which quantified the percentage of use of each tool was carried out. The result shows that the most frequently used tools are e-mail, telephone and instant messaging tools, in that order, either using Messenger or OCS [see Figure 2].

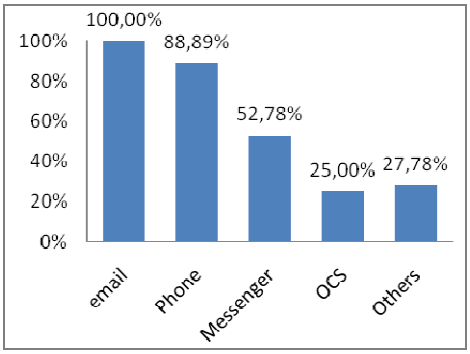

Figure 2: Communication tools

This analysis allowed us to conclude that it is highly appropriate that a tool that supports communication in DSD should have information concerning e-mails, and how to contact by telephone (telephone number) and the option for instant messaging.

We shall attempt to answer the question: *Do any significant differences exist in the use of the communication tools which might justify the difference in data items knowledge between the companies?* as follows.

In this case an analysis of each company was carried out in order to see whether any communication lack existed. This analysis provided us with very similar results. For example, upon comparing companies A and C [see Figures 3 and 4] it will be noted that company A uses slightly more synchronous tools than company C. But there were no significant differences between the companies.

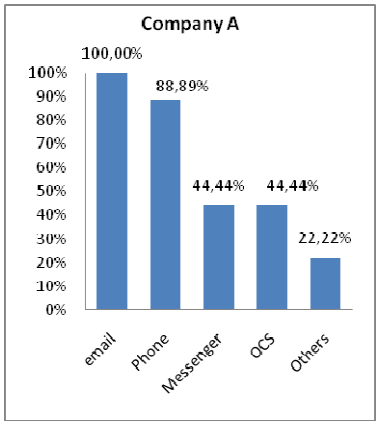

Figure 3: Company "A" Communication Tools

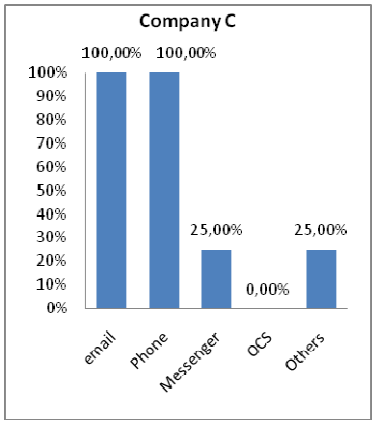

Figure 4: Company "C" Communication Tools

The study concerning the tools used by each company did not, therefore, obtain significant results. This is owing to the fact that the apparent lacks of the use of communication tools are not reflected, and this justifies the differences between the data items known by each of the companies in the study. What is demonstrated is the importance of companies that work in DSD using communication tools in a synchronous and an asynchronous manner.

In summary, this study allowed us to obtain a new requirement – the seventh one- with regard to the necessary communication tools. It could be stated as follows: The tool must:

- Provide an asynchronous messaging system along with information concerning the company's e-mail.
- Provide telephone contact information (telephone number).

- Provide a synchronous messaging service.

The following section explains how the tool that we have developed (Trusty) solves each of the requirements detected.

4 Trusty

The tool denominated as Trusty is a system based on the concept of social networking and has been designed for companies that work with a distributed software development model. The principal objective of the tool will therefore be to attempt to minimize problems caused by the characteristics of global development.

We have principally concentrated on the problems caused by the lack of trust, lack of information between colleagues, and communication, coordination and control problems.

The users of the tool will be members, or affiliate members, of the same company, who will be located in different countries or even different continents.

It is common that the information that is known about a work colleague is not sufficient to permit the effective development of work activities. This may be for two reasons:

- The user does not have access to this information.
- The user does not consider this information to be useful and does not, therefore, know it.

In this case, the tool offers information of both a work-related and a personal nature, which favours personal and human dealings between co-workers. Nevertheless, the users will always be able to determine which information they wish to share and with whom to share it, thus maintaining their intimacy and privacy.

This is based on the idea that if people have more information about each other it is possible for greater trust to grow between them, thus improving relationships and communication between co-workers [Panteli and Tucker 2009].

In our previous research study [Aranda et al. 2010] the data were classified into three groups according to their importance, and the users' information is therefore shown via three profiles:

- *Public Profile* [see Figure 5]: this can be accessed by all members of the company. This profile shows all the information which is necessary to establish communication with a work colleague. The members' photographs appear, since this is also considered to be important [Rusman et al. 2010]. It is also possible to view the members' work hours and the time at which they prefer to be contacted.
- *Project group profile* [see Figure 6]: this profile can be accessed by all the members of a project. This profile shows all the data which is relevant to the project (members' roles in the project, tasks upon which they are working) and any other useful information, such as the programming languages and technologies in which they are proficient, etc.
- *Personal profile* [see Figure 7]: this profile can only be accessed by work colleagues with whom a friendship exists. This profile contains data of a

personal nature, such as contact information of a more personal nature, culture, education, interests and hobbies.

As with the study explained in the previous section, a total of 7 requirements were obtained. An explanation of how these were decided upon will be provided as follows. Figures 5, 6 and 7 show the various user profiles.

In order to indicate where each of the pieces of data detected in the requirements is reflected, each figure has been marked with the corresponding data item's identifier through the use of a circle. For example, in Figure 5 the data item *"Other person's nickname"* with the identifier 3 is marked with an orange circle with the number 3 next to the user's nickname, which in this case is "Thomas".

In the first place, the role study determined that there are different information needs, according to the role played. The tool resolves the information gap, and offers the different profiles the needs found as it is following explained:

The *first requirement* is to offer those data items that *programmers* consider to be useful but do not know. Among these we have discovered:

- "*Technologies in which s/he has experience*" (41): this data item has been added to the "Project group profile" [Figure 6]. In this field, the user can introduce the technologies in which s/he has working experience, and which may be useful for other project colleagues to know.
- "*Programming languages in which s/he is an expert*" (42): this data item has also been added to the "Project group profile" [Figure 6]. It shows the programming languages with which the user has had working experience, and can be shared with project colleagues.
- "*Languages spoken*" (15): this can be found in the "Public Profile" [Figure 5], and provides a list of all the languages that the user knows at a spoken, read and written level. It is considered important to know what other languages a person has knowledge of in order to attempt to maintain smooth communication and reduce misunderstandings, particularly between users whose maternal languages are different.
- "*Planned meetings (date, time and place)*" (35): The system has a section for events in which it is possible to create and consult general work events, group project events, and even personal events. [Figure 5].
- "*Level of English*" (16): This item is located in the "Public Profile" [Figure 5]. English is represented amongst the languages known by the user, at its spoken, read and written level, as explained above. The tool also indicates the native language, this information can be useful, for instance, in the case that two people share the same native language and want to use it to clarify misunderstandings.

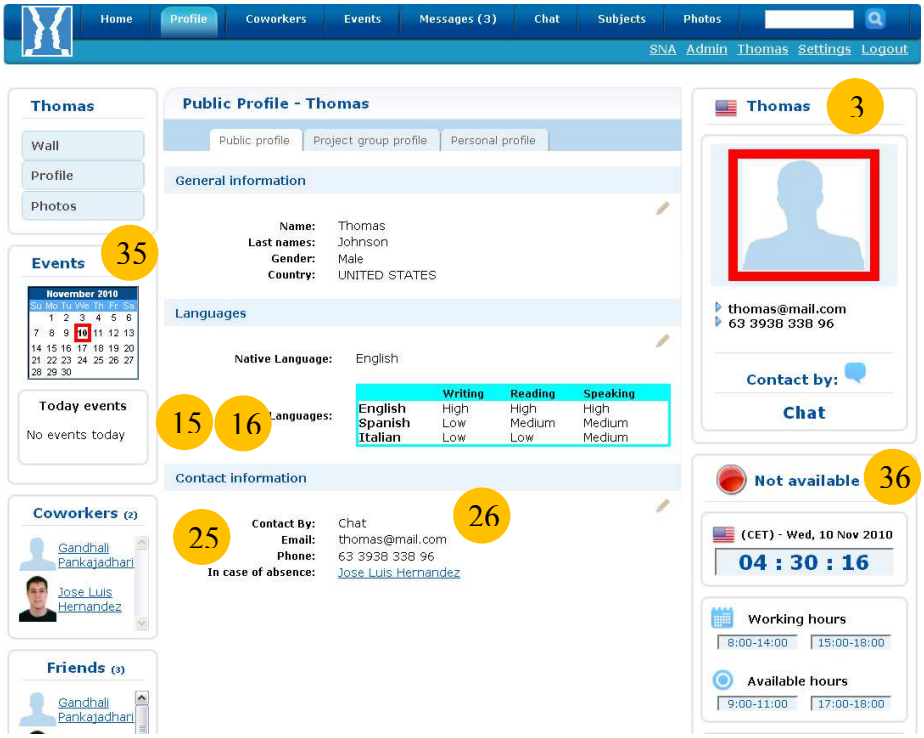

Figure 5:(Trusty) Public Profile

The *second requirement* is, in this case, to offer *analysts* the following data items which they do not know and are interested in obtaining:

- "*Programming languages in which s/he is an expert*" (42): this item has been added to the "Project group profile" [Figure 6]. Here it is possible to specify the programming languages with which the users have experience.
- "*Other person's stress levels*" (14): this has not yet been located in any specific section of the tool, and it is considered as future work.

The *third requirement* found is to offer the *project managers* those data items of which they do not have sufficient knowledge but consider important. Amongst these we can find:

- "*Other person's availability*" (36): User availability is a data item that was considered to be important after carrying out the study, and has therefore been situated in the "Public profile" [Figure 5]. The users' states are calculated automatically and according to their work hours and availability. The latter is the period of time within work hours at which users prefer to be contacted. The states are represented by colours.
    - Red: Not available (Outside work hours).
    - Orange (Possibly not available (Within work hours but outside time at which available).

- Yellow: Available, but for short amount of time (Close to end of preferred availability time).
- Green: Available (Within work hours and within time at which available).

- "Preference in carrying out determinate role" (23): This data item is particularly interesting for project managers, since each user can decide his/her preferred work role in accordance with his/her working experience. This data item is reflected in the "Project Group Profile" [Figure 6].

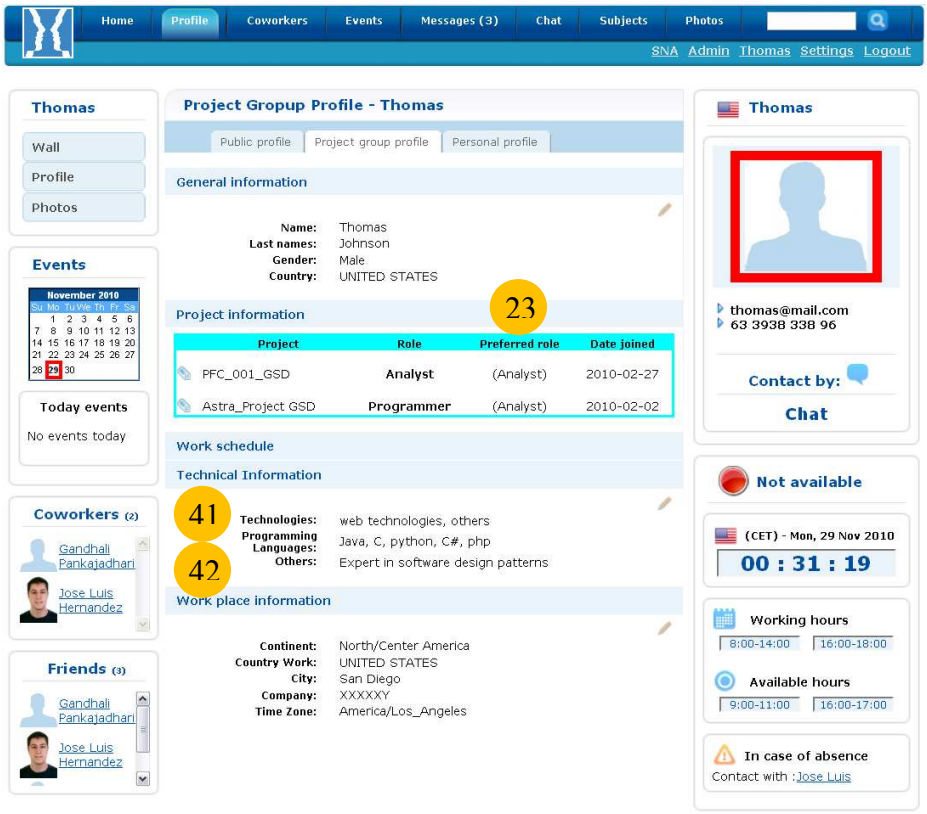

Figure 6: (Trusty) Project Group Profile

The data shown in Table 12, which have been obtained by analysing the project manager's answers, are considered the third requirement for the tool. However, they have not been reflected in any specific profile, but will be dealt with in future work.

Data Item
Other person's stress levels
Other person's emotional state

Table 12: Data item which are considered important by project managers but are not reflected in this version of the tool.

Having dealt with the analysis requirements by roles, it is time to tackle the requirements that resulted from the data items study regarding usefulness in communication, coordination and control.

We shall commence with those data that were considered to be useful for *communication*, but were not well-known. Upon providing the following data the *fourth requirement* will be revolved:

- "*Other person's nickname*" (3): This data item, which appears in the "Public Profile" [Figure 5], is a friendly means of addressing the users, thus creating confidence, since a person usually feels more relaxed when addressed with a nickname with which s/he feels comfortable.
- "*Cultural aspects of other person's country*" (9): This data item appears in each user's "Personal Profile" [Figure 7]. The users can introduce the principal cultural aspects of their countries into this field, including customs or religious beliefs.
- "*Other person's availability*" (36): The state of availability shown in the "Public Profile" [Figure 5].
- "*Age*" (5): The user's age is considered to be of a personal nature and is therefore located in the "Personal Profile" [Figure 7].
- "*Personal e-mail address*" (27): This item is located in the "Personal Profile" [Figure 7] and contains information of a more personal nature with regard to communication.
- "*Birthday*" (8): The date of someone's birthday is also a piece of data of a more personal nature, but the respondents considered it to be important for communication. It is therefore located in the "Personal Profile" [Figure 7].
- "*Companies in which s/he has worked*" (20): Those data items related to work experience and curriculum are located in the user's "Personal Profile" so that they can be viewed by those colleagues with whom the user has a better personal relationship.

Information requirements also appear in the case of *coordination,* of which we can highlight:

- "*Other person's availability*" (36): The other person's state of availability is located in the "Public Profile" as was mentioned previously [Figure 5].
- "*Years of work experience*" (21): This data item is located in the "Personal Profile" [Figure 7], which indicates how many years of work experience the users have had, along with other data such as the companies for which they have worked.

- "*Preference in carrying out determinate role*" (23): This is shown in the "Project Group Profile" and is principally of interest to project managers, as was mentioned earlier.

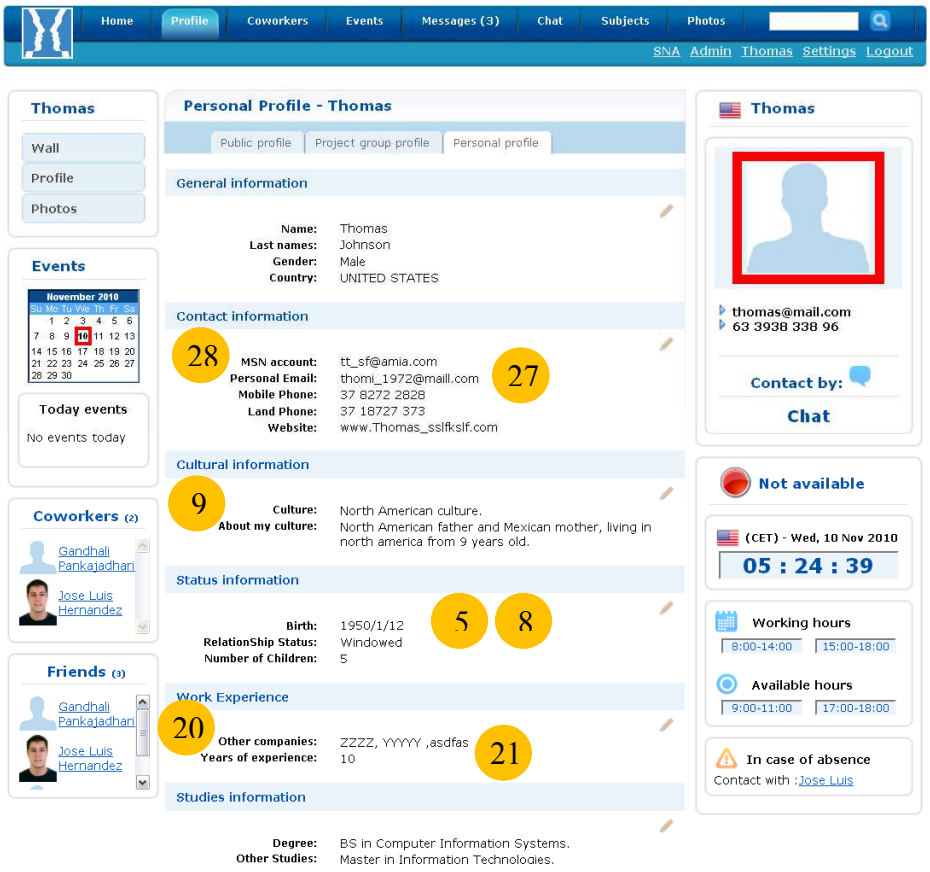

Figure 7: (Trusty) Personal Profile

The *sixth requirement* is to offer data items which are considered to be useful for *control*, and are not well-known. The data item that will be of particular interest to project managers is the following:

- "*Years of work experience*" (21): This item is located in the "Personal Profile" [Figure 7].

The data items listed in Table 13 are those which are considered to be useful in communication, coordination and control analyses, and are not known. They have not been considered in this version of the tool, and will be future work.

Data Item
Other person's stress levels
Other person's emotional state
Other person's stress levels

Table 13: Data items considered to be important for communication, coordination and control which are not known and are not considered in this version of the tool.

Finally, the *seventh requirement* which arose after the analysis of *communication tools* will be soved. The tool offers systems for synchronous and asynchronous communication, along with information concerning how to make contact.

It specifically implements the following communication systems:

- Asynchronous messaging systems: The tool implements a messaging system which is capable of attaching small documents.
- Synchronous messaging system (chat): The tool also implements a chat system in which various "rooms" of a general nature are created for the projects on which users are working, and with the possibility of sending private messages.
- Wall posters: Each user will have a wall on which the other users can leave public messages.
- Event management: An event system will be available in which the users will be able to create work-related events such as milestones or meetings.

Other means of contact are also offered:

- "*Contact phone number*" (25): Information regarding the contact 'phone number, which can be viewed in the "Public profile" [Figure 5].
- "*Business e-mail address*" (26): Information regarding the company's e-mail address. This information is also located in the "Public Profile" [Figure 5].
- Instant messaging: information concerning those instant messaging accounts which are most frequently used, such as "Messenger account" (28). This information is located in the "Personal Profile" [Figure 7], since the tool already provides a specific chat system.

The tool also provides the possibility of configuring the means of communication by which the user prefers to be contacted, thus controlling that the person is disturbed as little as possible during working hours. Similarly, in the case of unavailability, it is possible to select another person who can be contacted.

Moreover, in order to improve knowledge management, the system is synchronized with the project management programme, specifically the Ms Project system. This is useful in gathering the users' work-related information and in following the projects and tasks upon which they are working. It also permits the creation of work-related events, providing agenda functionality for deadlines, task completion, meetings and project milestones.

5 Conclusions and Future Work

This study has demonstrated that users who depend on their work functions have different information needs. The objective is for each user to have the necessary information about work colleagues. Achieving this could have positive repercussions on the communication, coordination and control processes of projects, and raise the level of confidence.

Certain data items appear in various requirements and are not offered in this version of the tool. Of these we can mention stress levels, emotional state and personality [see Table 13]. The means by which we intend to incorporate these characteristics are described as follows.

- "*Other person's personality*": One way in which to characterize the user's personality is through his/her use and interaction with the tool. To do interaction analyses in social networks will be used. Therefore, it is possible to discover how many friends that person has, the messages sent, the events created or the profiles created. It would thus be possible to gain insight into the user's personality (extrovert, shy, a leader …).
- "*Other person's emotional state*": The user's emotional state is more difficult to reflect. But the users could, for example, configure their state in their profiles by including emotions and phrases that reflect their emotional states.
- "*Other person's stress levels*": Psycho-physiological studies concerning the use of peripheral temperature to measure user stress exist with regard to user stress levels. [Ruvalcaba et al. 1998]. We are currently in contact with a group from the University of Cólima (Mexico), which is working in this area, in order to study the possibility of collaboration in this line.

Before concluding, it is important to note that the requirements obtained from this study may be also useful for companies or labs which are interested in developing tools to support DSD. Therefore, this study might be one step towards improving the current tools and reducing the distance between engineers' needs and the services that current tools offer.

Acknowledgements

This work has been funded by ORIGIN Integrated Project (Intelligent Global Innovative Organizations) (IDI-2010043 (1-5)) funded by the Center for Industrial Technological Development (CDTI) and the European Regional Development Fund (FEDER). This work has been funded too by the PEGASO/MAGO project (Ministerio de Ciencia e Innovación MICINN and Fondos FEDER, TIN2009-13718-C02-01). It is also supported by ENGLOBAS (PII2I09-0147-8235), funded by Consejería de Educación y Ciencia (Junta de Comunidades de Castilla-La Mancha).

We would like to thank those companies which took part in this survey, which include INDRA and Alhambra EIDOS.

References

[Aranda et al. 2010] Aranda, G. N., Vizcaíno, A., Palacio, R. R. and Moran, A. L.: "*What Information Would You Like to Know about Your Co-worker? A Case Study";* (2010),

[Bos et al. 2002] Bos, N., Olson, J., Gergle, D., Olson, G. and Wright, Z.: "*Effects of four computer-mediated communications channels on trust development";* (2002). Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves. Minneapolis, Minnesota, USA, ACM.

[Carmel (1999)] Carmel, E.: "*Global Software Teams: Collaborating Across Borders and Time Zones"* (1999), Prentice Hall.

[Conchuir 2010] Conchuir, E. O.: "*Global software development: a multiple-case study of the realisation of the benefits";* (2010), University of Limerick, Department of Computer Science & Information Systems.

[Conchúir et al. 2009] Conchúir, E. Ó., Ågerfalk, P. J., Olsson, H. H. and Fitzgerald, B. "*Global software development: where are the benefits?*" (2009) Commun. ACM 52(8): 127-131.

[Corbitt et al. 2004] Corbitt, G., Gardiner, L. R. and Wright, L. K.: "*A Comparison of Team Developmental Stages, Trust and Performance for Virtual versus Face-to-Face Teams";* (2004),

[Danis 2000] Danis, C. M. "*Extending the concept of awareness to include static and dynamic person information*" (2000) SIGGROUP Bull. 21(3): 59-62.

[Espinosa 2004] Espinosa, J. A.: "*The Effect of Time Separation on Coordination Costs in Global Software Teams: A Dyad Model";* (2004),

[Jarvenpaa et al. 2004] Jarvenpaa, S. L., Shaw, T. R. and Staples, D. S. "*Toward Contextualized Theories of Trust: The Role of Trust in Global Virtual Teams*" (2004) Info. Sys. Research 15(3): 250-267.

[Oshri et al. 2007] Oshri, I., Kotlarsky, J. and Willcocks, L. P. "*Global software development: Exploring socialization and face-to-face meetings in distributed strategic projects*" (2007) The Journal of Strategic Information Systems 16(1): 25-49.

[Panteli and Tucker 2009] Panteli, N. and Tucker, R. "*Power and trust in global virtual teams*" (2009) Commun. ACM 52(12): 113-115.

[Rusman et al. 2010] Rusman, E., Van Bruggen, J., Sloep, P., Valcke, M. and Koper, R.: "*Can I trust you? - Personal profiling for a first impression of trustworthiness in virtual project teams";* (2010).

[Ruvalcaba et al. 1998] Ruvalcaba, S. P., Mascorro, G. E. and Trejo, B. D. "*Cambios en los Correlatos Psicofisiológicos de Temperatura Periférica a lo largo de 14 Talleres de Manejo de Estrés*" (1998) UNAM.

[Shachaf 2008] Shachaf, P. "*Cultural diversity and information and communication technology impacts on global virtual teams: An exploratory study*" (2008) Inf. Manage. 45(2): 131-142.

[Walther 2005] Walther, J. B.: "*The Rules of Virtual Groups"*. 38th Hawai International Conference on System Sciences, Hawai (2005),

Appendix 1

Data items included in the questionnaires.

ID	Data item
01	First name
02	Last name
03	Nickname
04	Gender
05	Age
06	Marital status
07	Number of children
08	Birthday
09	Cultural aspects related to his/her country
10	Religion
11	Personal interests
12	Personality
13	Emotional state
14	Stress levels
15	Languages s/he speaks
16	Level of knowledge of English
17	Major degree
18	Year in which his/her degree was completed
19	University at which the degree was obtained
20	Companies in which s/he has worked
21	Years of work experience
22	Role in project
23	Preference in carrying out determinate role
24	Time difference with the city where he/she works
25	Contact phone number
26	Business e-mail address
27	Personal e-mail address
28	Messenger account (or Office Communications Server (OCS))
29	Personal Facebook (or other social network) account
30	His/her working hours
31	Time when s/he prefers to be contacted
32	Means by which other person prefers to be contacted (e-mail, telephone, etc)
33	List of public holidays in other person's location
34	Applied for and approved holidays
35	Planned meetings (date, time and place)
36	His/her availability at a given moment (available, having breakfast, in meeting, time inactive, etc.)
37	Tasks on which s/he is working at that moment and possible future tasks
38	Date s/he joined project
39	Director's (or Supervisor's) Name and Contact Information
40	Person to contact in case of absence
41	Technologies in which s/he has experience
42	Programming languages in which s/he is an expert
43	I usually know the people I work with in a distributed or global manner personally

Table 14: Data items included in the questionnaires