2010 IEEE International Conference on GLOBAL SOFTWARE ENGINEERING

23-26 August 2010 • Princeton, New Jersey, USA

Sponsored by

IEEE

Published by

IEEE computer society

CPS Conference Publishing Services

IEEE Computer Society Order Number P4122
ISBN-13: 978-0-7695-4122-8
BMS Part # CFP10ICG-PRT
Library of Congress Number 2010927211

Additional copies may be ordered from:

IEEE Computer Society
Customer Service Center
10662 Los Vaqueros Circle
P.O. Box 3014
Los Alamitos, CA 90720-1314
Tel: + 1 800 272 6657
Fax: + 1 714 821 4641
http://computer.org/cspress
csbooks@computer.org

IEEE Service Center
445 Hoes Lane
P.O. Box 1331
Piscataway, NJ 08855-1331
Tel: + 1 732 981 0060
Fax: + 1 732 981 9667
http://shop.ieee.org/store/
customer-service@ieee.org

IEEE Computer Society
Asia/Pacific Office
Watanabe Bldg., 1-4-2
Minami-Aoyama
Minato-ku, Tokyo 107-0062
JAPAN
Tel: + 81 3 3408 3118
Fax: + 81 3 3408 3553
tokyo.ofc@computer.org

Individual paper REPRINTS may be ordered at: <reprints@computer.org>

Editorial production by Patrick Kellenberger
Cover art production by Mark Bartosik
Printed in the United Stated of America by Applied Digital Imaging

IEEE Computer Society
***Conference Publishing Services* (CPS)**
http://www.computer.org/cps

Welcome Message
ICGSE 2010

The 5th International Conference on Global Software Engineering (ICGSE) brings together researchers and practitioners interested in exploring how globally distributed teams work and how the challenges posed by global software engineering can be met. This conference is a forum at the intersection of software engineering, communications, collaboration, business, and cultural aspects that influence human behaviors when faced with software development in global environments. This is the fifth in an annual series of international conferences on Global Software Engineering that began almost five years ago in Florianópolis, Brazil, in 2006. In our subsequent sessions we have learned from different cultures, where global software engineering is thriving: Germany, India, Ireland, and in our fifth edition: North-America.

The research agenda presented at the previous meetings has influenced the research field in global software engineering. We notice more formal modeling and empirical studies, and increased focus on the education of the new software engineers.

This year's technical program is as strong as ever. We received 55 submissions which were thoroughly reviewed by three or four reviewers from an expert program committee. From these 23 were accepted as research papers, 4 as industrial experience papers and 4 as educational papers. The Conference is structured in 9 sessions over 3 days in a single track. Our technical program attempts to address this broad area by presenting new insights into new tools, management, processes, human aspects, and teaching as applied to global software engineering.

The conference provides great opportunities for open discussion of issues and research directions, prompted by three workshops (PARIS'10, Knowing, REMIDI) and one panel. Our panel topic should provide an opportunity for us to travel virtually into the future and to speculate on how the state of the art in communication technologies will likely impact the workplace, specifically with the modern advances in communication technologies.

We offer our sincere thanks to the many individuals and organizations that helped make this year's Conference possible: the IEEE Computer Society, Siemens AG, Siemens Corporate Research, the ICGSE steering committee, the ICGSE 2010 program committee, and the ICGSE 2010 organizing committee. Finally, we would like to thank the authors, the tutorial and keynote speakers, and the workshop organizers and participants for making this year's ICGSE an exciting event.

Welcome to ICGSE 2010 in Princeton and enjoy the Fifth International Conference on Global Software Engineering!

August 2010

Alberto Avritzer, General Chair
Yael Dubinsky, Program Co-chair
Allen Milewski, Program Co-chairs

2010 International Conference on Global Software Engineering

ICGSE 2010

Table of Contents

Tools I: Support and Use

Processes and Practices

Management Environments I

Industry Experiences

Human Aspects

Preparing Software Engineers for Global Software Environments

Management Environments II: Assessing the Risk

Tools II: Emerging Tools

Management Environments III: Understanding Effort in a Global Context

PARIS'10: Methods and Tools for Project/Architecture/Risk Management in Globally Distributed Software Development Projects

REMIDI'10: Fourth International Workshop on Toll Support Development and Management in Distributed Software Projects

KNOWING'10: Knowledge Engineering in Global Software Development

Doctoral Symposium

What information would you like to know about your co-worker? A case study

Gabriela N. Aranda
GIISCo Research Group
Computing Sciences Department
Universidad Nacional del Comahue
Neuquén, Argentina
garanda@uncoma.edu.ar

Aurora Vizcaíno
ALARCOS Research Group
Information Systems and Technologies Department
UCLM-INDRA Research and Development Institute
Universidad de Castilla-La Mancha
Ciudad Real, Spain
aurora.vizcaino@uclm.es

Ramón R. Palacio[1] and
Alberto L. Morán[2]
[1] Facultad de Ingeniería
[2] Facultad de Ciencias
Universidad Autónoma de Baja California
Ensenada, México
{rpalacio, alberto_moran}@uabc.mx

***Abstract*—The lack of face-to-face interaction in Distributed Software Development (DSD) projects is a great obstacle for informal communication, which is the way in which we normally discover more about the co-workers along with building trust. Since we are attempting to reinforce the team spirit and improve trust between co-workers, we have analyzed what information software engineers consider to be important to know about their virtual colleagues. In this paper we describe the preliminary results obtained from this study and outline the features of a tool, called Trusting Social Network, with which to support DSD and help stakeholders to know each other better.**

Keywords—Global software development; trust building; informal communication

I INTRODUCTION

It could be said of the last decade that distributed software development (DSD) has become a frequent means of developing software. The tendency during recent years shows that the number of enterprises whose teams are distributed in different cities, countries, and even continents, is continually growing. However, although distributed software development projects are more and more common, it is well known that they tend to suffer from both delays and a wide range of problems [8]. One of these problems is that the lack of face to face communication signifies that people on distant sites have very little informal communication [5]. According to observational studies [10], informal communication supports different kinds of functions such as work-related activity execution and coordination, along with social functions which are also related to trust building [4]. Therefore, considering that informal communication is a means to discover more about the people that collaborate in a team, and that this characteristic is quite limited in a DSD environment, we have focused our research on the information that people usually know and would like to know about their distributed co-workers. To do this, we have conducted a survey in different companies that run distributed software development projects in different cities in the same country (DSD) and also on sites that traverse countries' boundaries, i.e., a global software development (GSD) environment. By doing this, we have had the opportunity to analyze the survey data and consider the differences between both settings. The results of this survey are presented in this paper, along with a tool to support DSD by making interpersonal communication easier.

The remainder of the paper is structured as follows: First, an introduction to social networks is presented. We then introduce the research question behind our work and the survey carried out with people participating in virtual teams in order to discover what information they consider to be useful in DSD environments. We next present a prototype tool for social knowledge management in DSD projects, called Trusting Social Network (TSN), whose design was based on the results obtained in the survey previously presented. Conclusions and future work are addressed in the last section.

II BUILDING RELATIONSHIPS IN DSD

Since face-to-face meetings rarely occur in virtual teams, one of the problems that affects DSD is the loss of team spirit [1], which depends on how well people know each other and the level of trust between them. This lack of face-to-face communication signifies that improving communication and building trust in virtual teams necessitate the assistance of information and communication technologies [4]. For instance, web-based social networks (WBSN) may be a good mechanism through which to promote interactivity between individuals [9] and can be easily applied in virtual teams as a means to share information about individuals. However, using public social networks in a work environment may be a problem rather than a solution: the goal of public social networks is to share information with friends and family, concerning private aspects of life, rather than work related aspects. However, it is possible to develop a social network as an intra-organization tool, such as A.M.I.G.O.S [2], a WBSN for sharing information in a workplace, which is specially oriented towards improving knowledge management. A.M.I.G.O.S permits the creation of a profile for each team member (with information about the physical and e-mail addresses, languages, and a brief personal description), and it also promotes the contribution registration with failure or success stories [3]. In contrast to A.M.I.G.O.S, which is not specifically oriented towards distributed environments, we have focused our research on geographically distributed team characteristics, and

978-0-7695-4122-8/10 $26.00 © 2010 IEEE
DOI 10.1109/ICGSE.2010.24

particularly on the information that team members' profiles should include to provide their virtual co-workers with the most useful information. An explanation of how this goal has been addressed is provided in the following section.

III RESEARCH QUESTION

Since informal communication is the way in which people usually get to know each other and build relationships in the real world, and this type of communication is quite limited in geographically distributed environments, we have focused our work on the following research question:

- What information do stakeholders consider important to know about their partners?

To answer this main research question we have proposed the following sub-questions:

- What data about co-workers do stakeholders usually know?
- How useful are the data that stakeholders usually know about their co-workers?
- What data about co-workers would stakeholders like to know?

A. Questionnaire definition and application

In order to answer the aforementioned research questions, we defined a questionnaire which was sent to a series of DSD project members, from different hierarchical levels.

The structure of the questionnaire was as follows: First, some questions about the respondent's experience in distributed or global software development projects were included. Second, a table was presented describing information that a person may know about others, and the respondents were asked to indicate:

- Do you usually know this data item about the co-workers on the other site?

Yes	Y
No	N
No Response	NR

- How useful do you consider this data item?

Very Useful	VU
Useful	U
Normal	N
Not very useful	NV
Not at all useful	NU
No Response	NR

The items included in the questionnaire are listed in Table I. In order to easily identify the data items, an ID code has been assigned to each data item according to its order of appearance in the questionnaire.

TABLE I. 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

Figure 1 shows an example of the table as it appears in the questionnaire, and a possible answer from a stakeholder.

Information	I usually know it		I think this information is				
	Yes	No	Very Useful	Useful	Normal	Not very useful	Not useful at all
Other person's first name	×		×				
Other person's last name	×		×				
Other person's nickname		×					×

Figure 1. Example showing the three first questions and a possible answer.

The questionnaire was sent to virtual team members in

four companies that run DSD and GSD projects. Companies 1, 2 and 4 are located in Spain and run both DSD and GSD projects. The third company is located the in the USA and runs GSD projects. When stakeholders were currently working on a DSD project but had previous experience in GSD projects; they were asked to answer the survey by considering themselves in a global scenario. After some weeks, 23 questionnaires were received: 7 from company 1, 7 from company 2, 4 from company 3, and 5 from company 4. Furthermore, 9 of the 23 people that filled in the questionnaire had experience in GSD projects. An explanation of how the data gathered by means of these questionnaires was analyzed is provided below.

B. Data item classification

Our first task was to organize the information gathered by means of the questionnaire in order to obtain a list of the items that the subjects taking part in the survey considered to be most useful.

To do this we defined a *usefulness indicator* (UI), which represents the proportion of stakeholders that consider a data item Very useful (VU) or Useful (U), in relation to the number of stakeholders that ranked this item. The *UI* formula is defined as follows:

$$UI_i = (VU_i + U_i) / (VU_i + U_i + N_i + NVU_i + NU_i)$$

where:

- i indicates the data item identification,
- VU_i indicates the number of people that considered data i "Very Useful"
- U_i indicates the number of people that considered data i "Useful"
- N_i indicates the number of people that considered data i "Normal",
- NVU_i indicates the number of people that considered data i "Not Very Useful", and
- NU_i indicates the number of people that considered data i "Not at all Useful".

As will be noted, we have not taken into account situations in which stakeholders have not ranked the usefulness of a data item, which are the No Response (NR) answers. The UI indicator therefore considers the proportion of positive answers over the number of actual answers.

On the other hand, we defined a *knowledge indicator (KI)* which represents the proportion of stakeholders that said they usually know this piece of information (YES), in relation to the number of stakeholders that answered the question. Therefore, the *KI* formula is defined as follows:

$$KI_i = Y_i / (N_i + Y_i)$$

where:

- i indicates the data item identification,
- Y_i indicates the number of people that said they usually know data i, and
- N_i indicates the number of people that said they do not usually know data i.

As occurred in the calculation of UI, we did not take into account situations in which stakeholders did not answer whether or not they knew a particular piece of data, which are the No Response (NR) answers. As was explained previously, we therefore considered the proportion of positive answers over the number of real answers.

The data collected by means of the survey and the KI and UI indicators are shown in Table II. For example, the first row corresponds to data item 26, which represents the "business email address": 21 people said they usually know this piece of data, 20 people said it is very useful (VU), 2 people said it is useful (U), and 1 person considered it not useful (NU). Table II shows data items which are ordered according to the usefulness indicator (UI), from that which is considered to be the most useful (data item 26, business e-mail address, with UI=0.96) to that which is considered to be least useful (data item 08, birthday, with UI=0.00).

TABLE II. UI AND KI INDICATORS FOR DATA ITEMS

ID	Knowledge				Usefulness						
	Y	N	NR	KI	VU	U	N	NVU	NU	NR	UI
26	21	2	0	0,91	20	2	0	0	1	0	**0,96**
22	17	6	0	0,74	15	6	0	0	1	1	**0,95**
01	19	4	0	0,83	19	2	0	0	2	0	**0,91**
32	12	11	0	0,52	11	8	0	0	2	2	**0,90**
31	10	13	0	0,43	8	11	0	0	2	2	**0,90**
25	18	5	0	0,78	18	2	0	1	2	0	**0,87**
35	10	13	0	0,43	11	8	1	0	2	1	**0,86**
30	14	9	0	0,61	12	6	0	0	3	2	**0,86**
33	14	9	0	0,61	12	6	0	0	3	2	**0,86**
41	9	14	0	0,39	9	9	1	0	2	2	**0,86**
42	9	14	0	0,39	9	9	1	0	2	2	**0,86**
34	8	15	0	0,35	7	10	1	0	3	2	**0,81**
40	6	17	0	0,26	6	11	1	0	3	2	**0,81**
36	8	15	0	0,35	7	9	1	0	4	2	**0,76**
24	15	8	0	0,65	12	5	0	0	6	0	**0,74**
37	12	11	0	0,52	11	4	0	0	6	2	**0,71**
43	5	13	5	0,28	5	7	1	0	4	6	**0,71**
28	12	10	1	0,55	13	1	2	1	3	3	**0,70**
15	12	11	0	0,52	11	4	0	2	5	1	**0,68**
23	5	13	5	0,28	5	6	1	3	2	6	**0,65**
39	11	12	0	0,48	9	3	1	2	4	4	**0,63**
02	14	9	0	0,61	11	3	4	0	5	0	**0,61**
16	10	13	0	0,43	9	5	2	2	5	0	**0,61**
38	7	16	0	0,30	3	10	0	3	6	1	**0,59**
14	5	13	5	0,28	5	5	1	2	4	6	**0,59**
12	4	14	5	0,22	4	5	2	2	4	6	**0,53**
21	6	17	0	0,26	3	6	2	4	7	1	**0,41**
09	5	18	0	0,22	2	7	2	5	6	1	**0,41**
13	3	15	5	0,17	3	3	3	1	7	6	**0,35**
17	5	18	0	0,22	3	4	4	5	5	2	**0,33**
04	7	16	0	0,30	2	5	1	11	4	0	**0,30**
27	5	18	0	0,22	5	1	5	8	3	1	**0,27**
03	3	20	0	0,13	3	3	4	6	6	1	**0,27**
11	2	21	0	0,09	0	4	3	10	4	2	**0,19**
05	1	22	0	0,04	1	3	5	9	4	1	**0,18**
20	2	21	0	0,09	1	2	3	6	10	1	**0,14**
06	0	23	0	0,00	0	2	2	18	0	1	**0,09**
07	0	23	0	0,00	0	1	3	16	2	1	**0,05**
10	0	23	0	0,00	0	1	2	15	4	1	**0,05**
18	0	23	0	0,00	0	1	7	11	3	1	**0,05**
19	0	23	0	0,00	0	0	8	11	3	1	**0,00**
29	0	23	0	0,00	0	0	5	13	2	3	**0,00**
08	0	23	0	0,00	0	0	3	14	5	1	**0,00**

Analyzing the KI and UI, from Table II, we have noticed that, for all the data items under study, the UI indicator is higher or equal to the KI indicator (Figure 2). That would indicate that, in general, people in GSD are interested in knowing more about their co-workers, except for a few data items that have obtained value 0 for both indicators (like birthday).

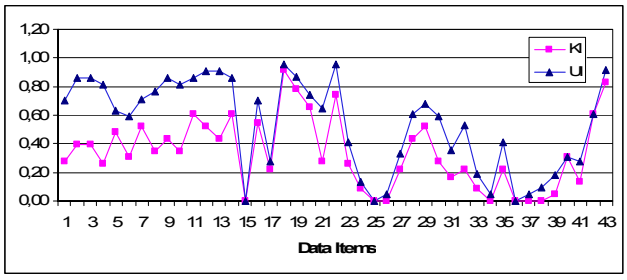

Figure 2. Analyzing the KI and UI indicators.

Having this in mind, and based on the results shown in Table II, we propose to define a model to classify the data items. This model will be useful in determining what information must be added and what information can be ignored in a tool to support DSD. Moreover, the results obtained by means of the analysis of the previous questionnaire were the basis for the requirements of the tool we are developing. Thus, in order to classify the data items in our list, we have defined three levels for the UI indicator and two levels for the KI indicator. First, for the UI indicator, we have defined the labels A, B, and C as follows:

- A: data items with a high UI (UI >= 0.66)
- B: data items with medium UI (0.33 < UI < 0.66)
- C: data items with low UI (UI <= 0.33)

Similarly, in the case of the KI indicator we have defined two labels 1 and 2 as follows:

- 1: data items with high KI (KI >= 0.50)
- 2: data items with low KI (KI < 0.50)

Therefore, by combining both sets of labels, we have defined a matrix (shown in Figure 3), whose intersections define the categories for the data items under study.

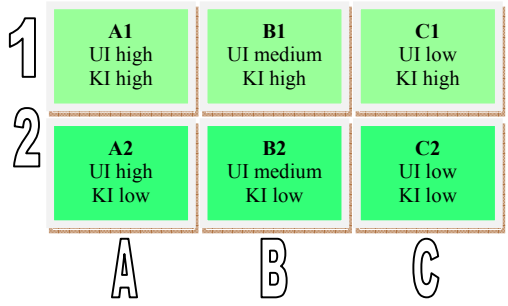

Figure 3. Matrix for data item classification.

Based on this matrix, we have defined the data item classification model as follows:

Category A1 (UI high – KI high): people in DSD projects usually know these data items and consider them to be very useful for distributed work. We consider that tools for supporting DSD should help people to share these data items and always be "close at hand". The data items that fit in this category are listed in Table III.

TABLE III. DATA ITEMS IN CATEGORY A1

ID	Data item description	KI	UI
26	Business e-mail address	0,91	0,96
22	Role in project	0,74	0,95
01	First name	0,83	0,91
32	Means by which other person prefers to be contacted (e-mail, telephone, etc)	0,52	0,90
25	Contact phone number	0,78	0,87
30	His/her working hours	0,61	0,86
33	List of public holidays in other person's location	0,61	0,86
24	Time difference with the city where s/he works	0,65	0,74
37	Tasks on which s/he is working at that moment and possible future tasks	0,52	0,71
28	Messenger account (or Office Communications Server (OCS))	0,55	0,70
15	Languages s/he speaks	0,52	0,68

Category A2 (UI high – KI low): people in DSD projects sometimes know these data items and consider them to be very useful for distributed work. We consider that tools for supporting DSD should help people to share these data items and assure access to them. The data items that fit in this category are listed in Table IV.

TABLE IV. DATA ITEMS IN CATEGORY A2

ID	Data item description	KI	UI
31	Time when s/he prefers to be contacted	0,43	0,90
35	Planned meetings (date, time and place)	0,43	0,86
41	Technologies in which s/he has experience	0,39	0,86
42	Programming languages in which s/he is an expert	0,39	0,86
34	Applied for and approved holidays	0,35	0,81
40	Person to contact in case of absence	0,26	0,81
36	His/her availability at a given moment (available, having breakfast, in meeting, time inactive, etc.)	0,35	0,76

Category B1 (UI medium – KI high): people in DSD projects usually know these data and consider them quite useful for distributed work. We consider that tools for supporting DSD should help people to share these data items. However, as these data items have been determined as very useful by less than 66% of the subjects in the survey, we recommend that people should have the opportunity to decide whether or not they wish to share this information. The data items that fit in this category are listed in Table V.

TABLE V. DATA ITEMS IN CATEGORY B1

ID	Data item description	KI	UI
02	Last name	0,61	0,61

Category B2 (UI medium – KI low): people in DSD projects sometimes know these data and consider them quite useful for distributed work. We consider that tools for supporting DSD should help people to share these data items. However, as these data items have been determined as very useful by less than 66% of the subjects in the survey, we recommend that people should have the opportunity to decide whether or not they wish to share this information. The data items that fit in this category are listed in Table VI.

TABLE VI. DATA ITEMS IN CATEGORY B2

ID	Data item description	KI	UI
23	Preference in carrying out determinate role	0,28	0,65
39	Director's (or Supervisor's) Name and Contact Information	0,48	0,63
16	Level of knowledge of English	0,43	0,61
38	Date s/he joined project	0,30	0,59
14	Stress levels	0,28	0,59
12	Personality	0,22	0,53
21	Years of work experience	0,26	0,41
09	Cultural aspects related to his/her country	0,22	0,41
13	Emotional state	0,17	0,35
17	Major degree	0,22	0,33

Category C1 (UI low – KI high): people in DSD projects usually know these data but rarely consider them to be useful for distributed work. In this case we suggest ignoring them or using only those that are considered useful for practical issues, but according to our survey no data items fit this category.

Category C2 (UI low – KI low): people in DSD projects sometimes know these data and rarely consider them to be useful for distributed work. In this case we suggest ignoring them or using only those that are considered useful for practical issues. The data items that fit in this category are listed in Table VII.

TABLE VII. DATA ITEMS IN CATEGORY C2

ID	Data item description	KI	UI
04	Gender	0,30	0,30
27	Personal e-mail address	0,22	0,27
03	Nickname	0,13	0,27
11	Personal interests	0,09	0,19
05	Age	0,04	0,18
20	Companies in which s/he has worked	0,09	0,14
06	Marital status	0,00	0,09
07	Number of children	0,00	0,05
10	Religion	0,00	0,05
18	Year in which his/her degree was completed	0,00	0,05
08	Birthday	0,00	0,00
19	University at which the degree was obtained	0,00	0,00
29	Personal Facebook (or other social network) account	0,00	0,00

C. Data analysis regarding GSD and DSD features

In order to detect differences concerning the answers between people working in DSD and GSD projects that could affect our study, we have analyzed the information separately for both groups. This was done by defining a group called DSD and a group called GSD. The DSD group is formed of people who only have experience in distributed environments, relating sites in different cities of the same country (14 people), and the GSD group is formed of people who have experience in global environments, signifying that they work with people from at least one different country (9 people). A series of data items that may represent certain differences with regard to appreciations in both groups is presented as follows.

Firstly, we analyzed the data regarding technology. As was expected, in the case of the business e-mail address (data 26) there are no differences between both groups. As can be seen in Figure 4, almost all the people in the DSD and GSD groups considered the business email address to be *very useful* or *useful*. However, with regard to the information concerning the phone number (data 25), shown in Figure 5, we discovered that this appears to be more important for people working on a DSD project: 93% of people in DSD projects consider it *very useful* versus the 56% in GSD projects. This could be explained if we consider that a DSD environment usually shares the same time zone and the same native language, and quite often the same timetable, and the phone can therefore be used as a common tool for contact, while this rarely occurs in GSD environments.

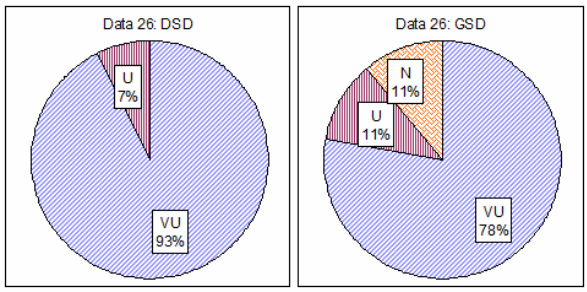

Figure 4. Comparison of "Business e-mail address" data

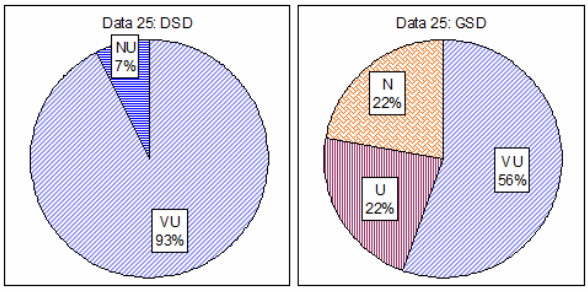

Figure 5. Comparison of "Phone number" data

Secondly, we have analyzed data items regarding language, bearing in mind that some notorious differences could exist between people in the DSD and GSD groups. For example, data item 15 (Languages spoken), was, as expected, higher for people in GSD projects: 62% of the stakeholders working in a GSD environment considered it *very useful*, while only 43% of people in DSD projects considered it *very useful*. However, as Figure 6 shows, less than 66% of the

people in the group with experience in GSD projects, considered it *useful* or *very useful*, which would indicate that, according to our model, this data item would correspond to category B. As this difference was unexpected, we analyzed the questionnaires for people working on GSD projects and realized that two Spanish people were working on GSD projects with people from countries in which the native language is also Spanish; they did not, therefore, consider it important to know about the other languages their co-workers speak. However, when we analyzed the same question without considering these two people, 83% of people in GSD considered knowing the languages that a person speaks to be *very useful* (Figure 7).

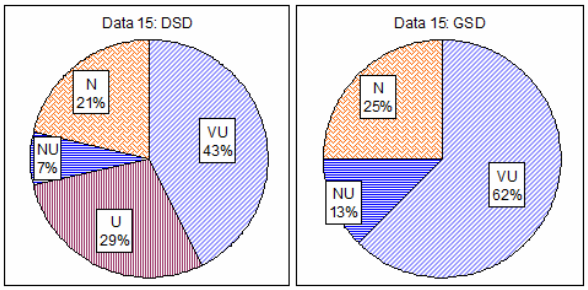

Figure 6. Comparison of "Languages spoken" data

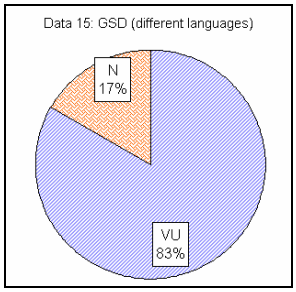

Figure 7. "Languages spoken" by people in GSD in countries with different native language.

Another piece of data regarding language is data item 16 (level of knowledge of English). In this case the differences between both groups are notorious (see Figure 8): 50% of people in DSD projects did not consider it useful to know their co-workers' level of English, while 78% of people in GSD projects considered it *very useful* or *useful*.

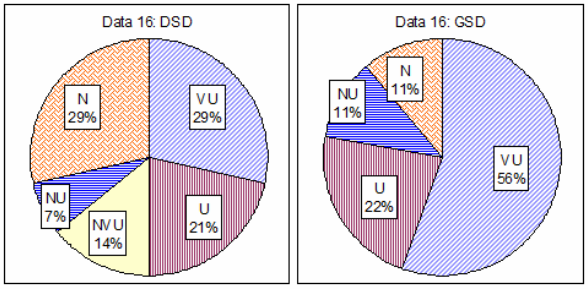

Figure 8. Comparison of "English level" data

This result was expected, since most of the participants in the DSD group were Spanish and shared their native language.

Finally we analyzed information about time difference and working hours, bearing in mind that important differences might exist between people in the DSD and GSD groups. As was expected, time difference (data 24) was considered to be much more important for stakeholders working in a GSD environment. This was in accordance with our previous expectations since participants working in DSD usually share the same time zone, or it is very close. As can be seen in Figure 9, in the group that has experience on GSD projects, 78% of them considered the information about time differences as *very useful* or *useful* while only 57% of people working on DSD projects considered it *very useful* or *useful*.

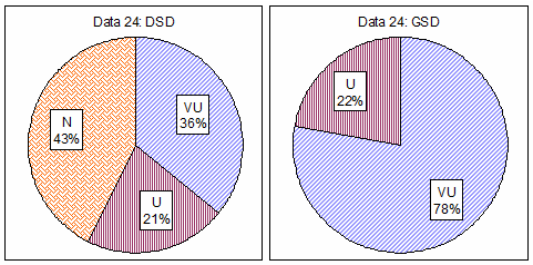

Figure 9. Comparison of "Time difference" data

Lastly, regarding the information on working hours (data item 30) we realized that it was considered in quite a similar manner by people working on both kinds of projects (Figure 10). For example, 58% of people in DSD considered it *very useful,* as did a similar percentage in GSD groups (57%), and the remaining people considered it *useful* or *normal*. Therefore, more than 66% of people in both groups considered these data items to be *very useful* or *normal*.

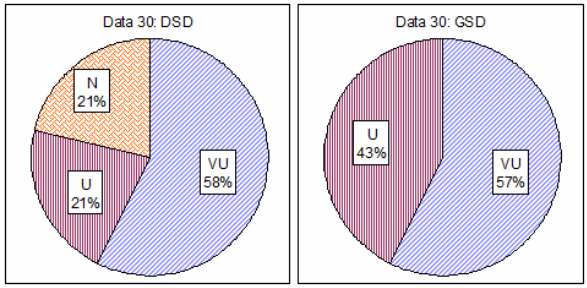

Figure 10. Comparison of "Working hours" data

IV TRUSTING SOCIAL NETWORK

As it was explained previously, we have developed this study in order to define the requirements for the tool Trusting Social Network (TSN), to support interpersonal communication in DSD projects. This tool has been designed to support the following capabilities, as is shown in Figure 11:

- To provide *useful information* about co-workers based on the features of distributed and global distributed environments.
- To provide tools to *support communication* in these environments
- To provide a mechanism for event creation and sharing, in order to *support coordination* between co-workers.
- To provide *reports and statistical analysis* about the social network supported by the tool.

Below we focus on the first goal, which is directly related to the survey and the classification of data items presented above.

Therefore, after attempting different designs and profiles, we decided that it would be most appropriate for TSN to enable users to have three different profiles in order to organize the information correctly:

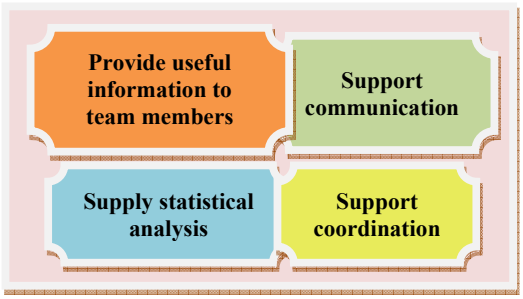

Figure 11. Main goals of TSN

- **Public profile**: This profile will show general information about stakeholders in the organization. It will be visible to all the people in the network, in order to provide a starting point for communication between all the people in the organization.
- **Project profile**: This profile will show project related information, such as current activities, role in the project, etc. It will be visible only to people taking part in the same project. The idea is to share information that might make communication and coordination easier when working on the same project
- **Personal profile**: This profile will show information specifically related to the person, such as his/her studies, working experience, etc. It will be visible only to people that have been previously accepted as "friends". Its goal is to give people the opportunity to share more information about themselves and to provide a channel for informal communication.

As is shown in Figure 12, these three profiles are nested; the public profile being that which provides less information but which can be accessed by everybody in the organization. On the other hand, the personal profile is that which shows more information but can only be accessed by a select group of people.

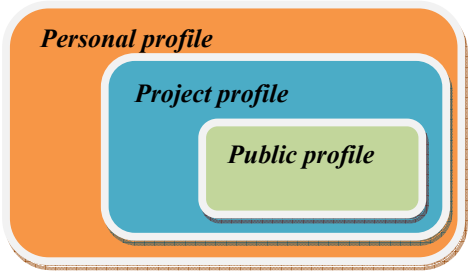

Figure 12. Relationship between profiles

The information in the public profile will also be visible in the project profile, while information in the project profile will be visible in the personal profile.

The main characteristics of these profiles are explained as follows.

A. Public profile

As was explained previously, this profile will be visible to everybody in the organization. Its goal is to provide: (1) information needed for identification; and (2) information needed to establish contact.

Thus, in order to identify a person, this profile includes the first and last name, gender, and the site where s/he works. We should underline that, according to our survey, the last name and gender were not considered useful data items; however, we decided to include them in the public profile because:

- A first name may be not sufficient to identify a person: more than one person on the same site may have the same first name. The last name may, therefore, be needed to identify a person.
- Globally distributed environments usually involve different cultures and sometimes gender cannot be easily inferred from the name, while it may be necessary for a first contact (for example, the appropriate use of the title Mr. or Ms.).

The second group of data items in this public profile includes contact information, such as the business e-mail address and the phone number. Furthermore, we have included the languages that a person knows, which can be used to contact him/her.

Space restrictions do not permit us to show a view of this profile, but it includes panel (1) and the first section of panel (3) in the project profile, which is shown in Figure 13.

B. Project profile

As we have already explained, the information in this profile will be visible to people working on the same project. The project profile will therefore include the basic information in the public profile (identification and contact information) and will add more detailed information that could be useful to the co-workers that need to make contact. For example, the project profile will include information

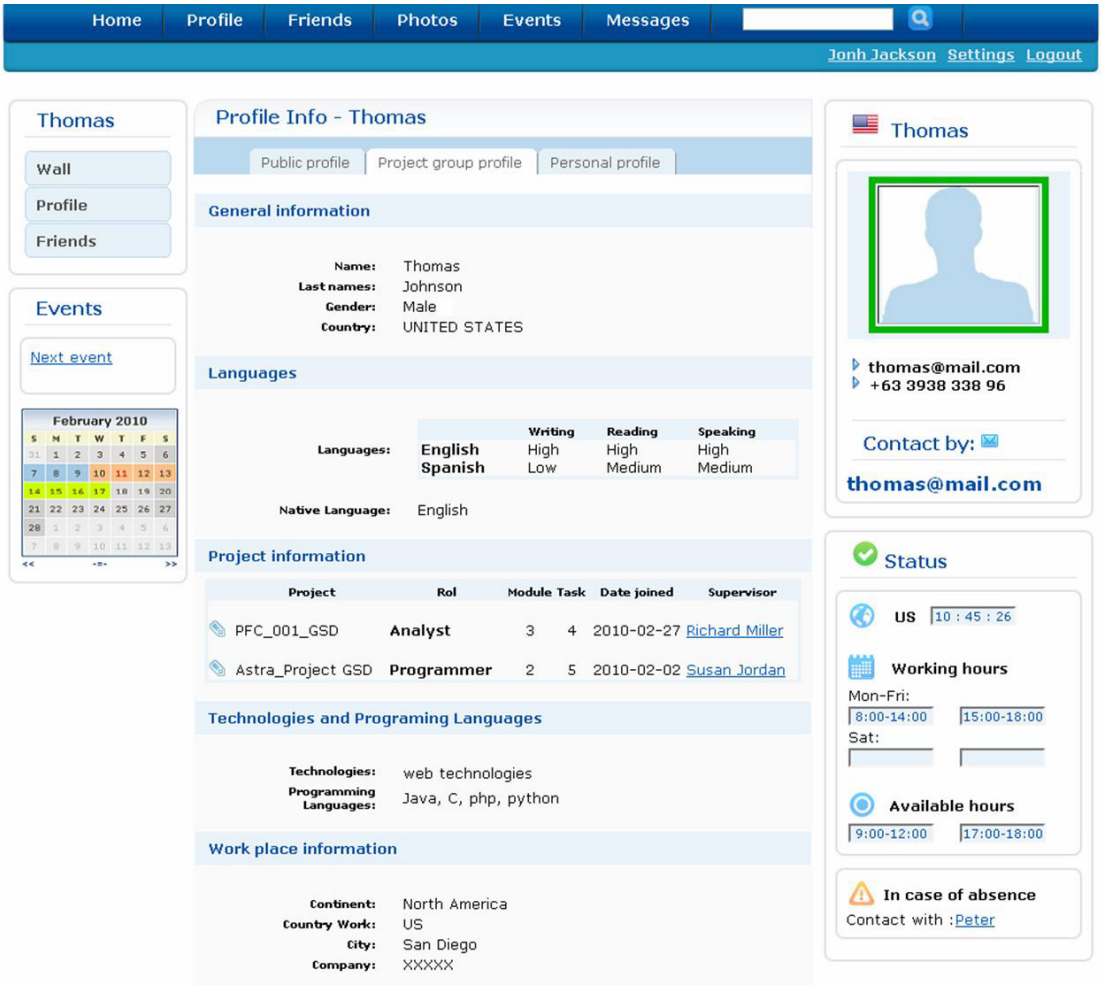

Figure 13. Project profile

about the different languages known, specifying the level for each one, so as to provide an opportunity to choose the language with which to make contact. This profile will also include information to help to choose the best moment to start communication: for example, it will include information about the working hours, the time at the site, the current status regarding availability, etc. Moreover, in order to make this information clearer for the users, we propose representing the user's status with a colour code similar to CWS [6]. This colour code is guided by the selective availability criterion [7], such as *"I am available only to people who are related to the task I am dealing with now and am not available to other people"*. To do this, we used different colours for the photo frame in panel (1) so as to indicate whether or not it is an appropriate moment to start a synchronous interaction with the other person. In Table VIII we present the colour code for this photo frame, taking into account the setting of the current status and the time at the site with regard to the hours at which that person prefers to be contacted.

TABLE VIII. PHOTO FRAME COLOUR CODE MEANING

Colour	Meaning	Status	Current time according to his/her preferred time to be contacted
Green	Available for contact	Available	In the interval
Yellow	Available for a short time	Available	Time is close to the end of his/her preferred interval
Orange	Probably available for contact	Available	Out of the interval
Red	Not available	Not available	---

The project profile will also include information about skills related to work such as knowledge of technologies or programming languages that may be useful for the project, and which is shown in panel (3).

An example of this profile is shown in Figure 13. In order to organize the information, in the top right-hand section, a small panel (1) shows a photo (if available), contact information (e-mail address and phone number), and the

preferred media through which to be contacted. In this case, according to the colour code meanings presented in Table VIII, the photo frame is green because the stakeholder's status is *available* and the time at the site is within his preferred time interval to be contacted. In a second panel (2), in the bottom right-hand section, information about current status, working hours and local time at his/her site is shown. The rest of the information is shown in the central panel (3), which is organized in different sections to facilitate access. Panel (4), in the left section, provides a quick access to the events, by means of a calendar that uses different colours for each type of event.

As was explained previously, part of the information in the project profile is shown in the public profile, including that of panel (1) and the "general information" of panel (3).

C. *Personal profile*

This information will be visible to those people that the profile's owner has chosen as "friends", which means it is a private profile. We have organized the information in this profile in a similar way to that of the project profile.

As can be seen in Figure 14, this profile also includes panel (1) in the top right-hand section, with the first name, and main contact information. In the case shown, the photo frame is orange because his status is *available* but the time at the site is not within the stakeholder's preferred time interval to be contacted. Panel (2) in the bottom right-hand section, also has information about the time at his site, current state, etc., while panel (4) in the left section, provides quick access to events information.

The rest of the information is displayed in panel (3), and is organized in different sections so as to facilitate access to them. This information includes mandatory data items in categories A and also the optional data items in category B. Moreover, if people wish, they can share photos, indicate their hobbies, and other information. They are therefore expected to feel more familiarized with their co-workers while the lack of trust that often takes place in DSD might decrease.

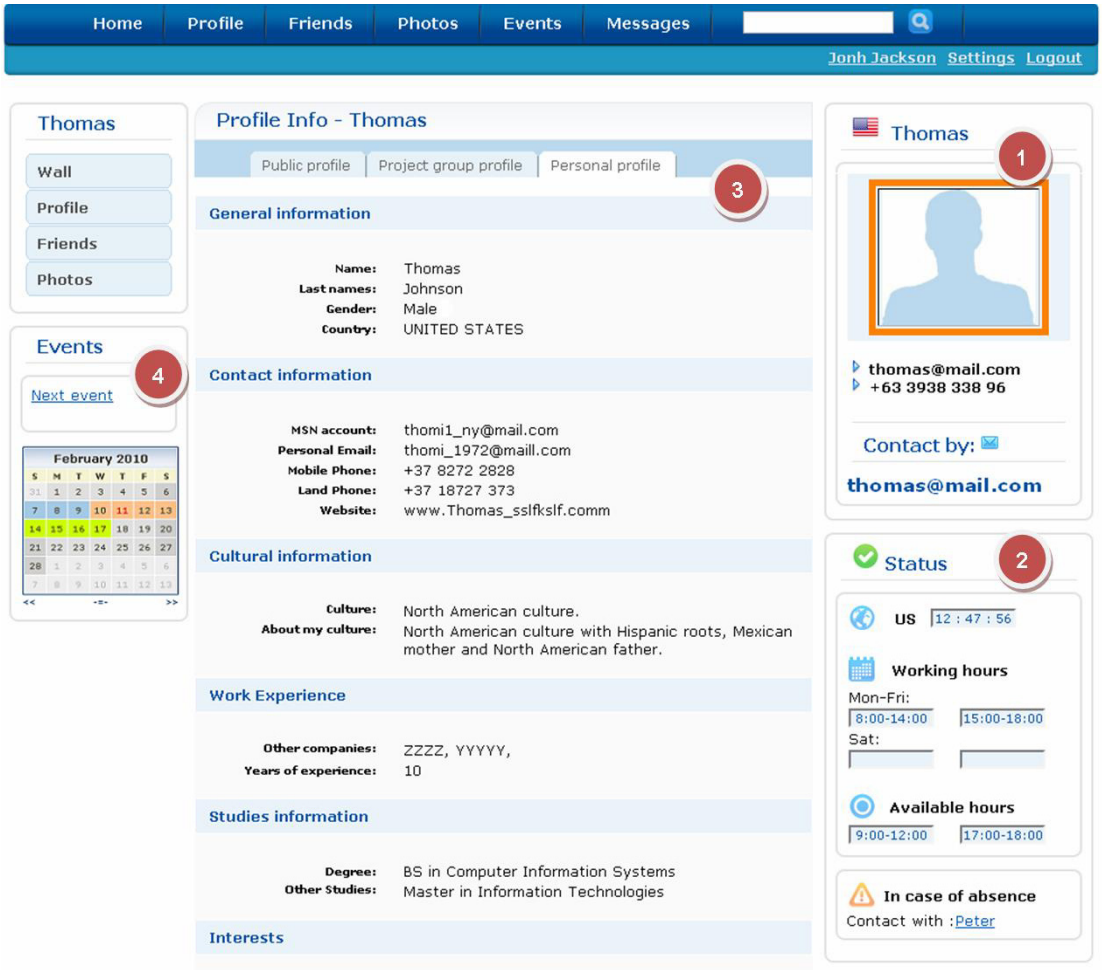

Figure 14. Personal profile

Summarizing, there are some characteristics that differentiate TSN from other social network tools (like Facebook, MySpace, LinkedIn, Orkut, etc.). For example:

- TSN allows defining three different user profiles (public, project and personal), while other tools define two levels (public and personal).
- TSN allows automatic generation of current status, by analyzing the users' preferences.
- TSN provides information not just about the user's current status but also about the time and the way s/he prefers to be contacted.

V CONCLUSIONS AND FUTURE WORK

Distributed software development has been definitively adopted by a growing number of organizations. However, the situation of teams being distributed throughout many distant sites has not only led to many benefits, but also to certain challenges. One of these challenges is that of generating the sense of team spirit and building trust between the members of virtual teams. In this paper we have presented the results of a survey that involved 23 people from DSD and GSD teams, and analyzed what information they considered to be most useful to know about their co-workers. We specially remark the importance of data items in subcategories A2 and B2, which is information rarely known in virtual environments but that is consider important for a wide number of the people in our survey. Such categories specially concern information about the best means and time to contact somebody as well as the person to contact in case of his/her absence. We think this found is important for people who develop DSD tools as added this information is not a bit issue but can improve people communication and coordination.

Furthermore, based on the results of this survey, we have obtained the requirements for a tool to support DSD projects with the goal of facilitating communication and coordination. To do this, TSN will use different mechanisms to indicate the best means and moment to contact a person. Moreover, the personal profile will enable co-workers to get to know each other better, thus increasing their team spirit and sense of trust. The project profile will also be useful to discover more about co-workers' qualification, their skills and the tasks they are working on at a particular moment. We are aware that the survey must be replicated in different organizations so as to validate the results we have obtained. Furthermore, as part of our future work, we plan analyzing the differences between people from different organizations and with different roles in the GSD project. Finally, our future work involves i) technical aspects, such as obtaining information about the context and the data flow of users' tasks; and ii) a validation of TSN through the use of a prototype by distributed software developers.

ACKNOWLEDGMENT

We would like to thank the people from the four companies that have filled in the questionnaire and contributed with their feedback. We would also like to thank José Luis Hernandez for helping us to analyze the data.

This work has been funded by the PEGASO/MAGO project (Ministerio de Ciencia e Innovación MICINN and Fondos FEDER, TIN2009-13718-C02-01). It is also supported by MEVALHE (HITO-09-126) and ENGLOBAS (PII2I09-0147-8235), funded by Consejería de Educación y Ciencia (Junta de Comunidades de Castilla-La Mancha), and co-funded by Fondos FEDER, as well as MELISA (PAC08-0142-3315), Junta de Comunidades de Castilla-La Mancha, Consejería de Educación y Ciencia in Spain. Also by the 04/E072 project de la Universidad Nacional del Comahue, from Argentina, and by UABC under grant 0207 of the XIII Convocatoria Interna de Proyectos de Investigación. The third author is supported by scholarship PROMEP/103.5/06/3244.

REFERENCES

[1] Carmel, E., *Global software teams: collaborating across borders and time zones*. Upper Saddle River, NJ, USA: Prentice Hall PTR. 269, 1999.

[2] Costa, R.A., Oliveira, R.Y.S., Silva, E.M., and Meira, S.R.L. "A.M.I.G.O.S: Knowledge Management and Social Networks". In *26th annual ACM International Conference on Design of Communication*. Lisbon, Portugal2008, pp.235-242.

[3] Costa, R.A., Silva, E.M., Neto, M.G., Delgado, D.B., Ribeiro, R.A., and Meira, S.R.L. "Social Knowledge Management in Practice: A Case Study". In *Groupware: Design, Implementation, and Use, 15th International Workshop, CRIWG 2009*. Peso da Régua, Douro, Portugal2009, pp.94-109.

[4] Harell, G. and Daim, T.U., "Virtual Teams and the Importance of Building Trust". *IT Professional*, 11(6): 2009, 46-49.

[5] Herbsleb, J.D. and Moitra, D., "Guest Editors' Introduction: Global Software Development". *IEEE Software*, 18(2): 2001, 16-20.

[6] Palacio, R.R., Morán, A.L., and González, V.M., "CWS: An Awareness Tool to Support Starting Collaboration in Global Software Development". *The Open Software Engineering Journal, Special Issue on Global Software Development and its Challenges*: 2010.

[7] Palacio, R.R., Morán, A.L., González, V.M., and Vizcaíno, A. "Collaborative Working Spheres as support for starting collaboration in distributed software development". In *3th International Conference on Computer Supported Cooperative Work in Design*2009, pp.636-641.

[8] Richardson, I., Casey, V., Zage, D., and Zage, W., Global Software Development – the Challenges. University of Limerick, Ball State University: SERC Technical Report 278. p. 10, 2005.

[9] Staab, S., Domingos, P., Mika, P., Golbeck, J., Ding, L., Finin, T., Joshi, A., Nowak, A., and Vallacher, R.R., "Social Networks Applied". *IEEE Intelligent Systems*, 20(1): 2005, 80-93.

[10] Whittaker, S., Frohlich, D., and Daly-Jones, O. "Informal workplace communication: what is it like and how might we support it?" In *SIGCHI conference on Human factors in computing systems: celebrating interdependence*. Boston, Massachusetts, United States1994, pp.131-137.