Analyzing Ontology as a Facilitator During Global Requirement Elicitation

Gabriela N. Aranda*

Aurora Vizcaíno*

Mario Piattini*

*Alarcos Research Group
Universidad de Castilla-La Mancha
Spain

+GIISCo Research Group
Universidad Nacional del Comahue
Argentina
• Motivation

• The framework

• Validation
Motivation

- Global Software Development (GSD) projects
  - Stakeholders are distributed throughout many geographically distanced sites

- Distance between members affects interpersonal communication
  - Lack of face-to-face communication
  - Time separation
  - Cultural diversity
    - Language
    - Customs

(Damian-Zowghi 2002, Richardson et al 2005)
Defining a framework for the Requirements Elicitation Process in GSD projects, considering stakeholders’ features in order to detect problems and propose solutions to decrease their influence and improve the process and the product quality.
### RE-GSD main phases

<table>
<thead>
<tr>
<th>Christel Framework</th>
<th>RE-GSD Framework</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preliminary Data Collection</td>
<td>Preliminary Data Collection</td>
</tr>
<tr>
<td>Requirements Gathering</td>
<td>Virtual team definition &amp; problem detection and solution</td>
</tr>
<tr>
<td>Rationalization and Evaluation</td>
<td></td>
</tr>
<tr>
<td>Prioritization</td>
<td></td>
</tr>
<tr>
<td>Integration &amp; Validation</td>
<td></td>
</tr>
</tbody>
</table>

(Christel 1992)
Phase 1: Preliminary Data Collection

1. Preliminary Data Collection

2. Virtual team definition & problem detection and solution

3. Requirements Gathering

- **Goal**: knowing as much as possible about the requirements elicitation scenario
### Phase 1: Preliminary data collection

#### Form 3A: Information about organization environment

<table>
<thead>
<tr>
<th>Question</th>
<th>Choices</th>
</tr>
</thead>
<tbody>
<tr>
<td>What kind of groupware tools are usually used in the organization?</td>
<td>e-mail, phone, instant messaging, discussion forum, other, share whiteboards, audio conference, video conference</td>
</tr>
<tr>
<td>Have you been trained to use them?</td>
<td>YES, NO</td>
</tr>
<tr>
<td>Do you consider you need to be trained to use any of them?</td>
<td>YES, NO</td>
</tr>
<tr>
<td>Which ones do you use more often and you know it better?</td>
<td>e-mail, phone, instant messaging, discussion forum, other, share whiteboards, audio conference, video conference</td>
</tr>
<tr>
<td>Which ones have not you used them before?</td>
<td>e-mail, phone, instant messaging, discussion forum, other, share whiteboards, audio conference, video conference</td>
</tr>
<tr>
<td>Why have not you used them before?</td>
<td>lack of knowledge, lack of infrastructure, organization policies, other</td>
</tr>
<tr>
<td>Which ones do you think that can improve your task?</td>
<td>e-mail, phone, instant messaging, discussion forum, other, share whiteboards, audio conference, video conference</td>
</tr>
<tr>
<td>What kind of tasks can be improved by them?</td>
<td></td>
</tr>
<tr>
<td>Do you know any organization policy that limits groupware tools at work?</td>
<td></td>
</tr>
<tr>
<td>When can they be applied</td>
<td></td>
</tr>
<tr>
<td>Which requirements elicitation techniques have you used before,</td>
<td></td>
</tr>
</tbody>
</table>
Requirements Engineering
“needs to be sensitive to how people perceive and understand the world around them, how they interact, and how the sociology of the workplace affects their actions”
(Nuseibeh, 2000)
Phase 1: Preliminary data collection

- Felder and Silverman Model
- Learning: Perception and Processing
- 4 categories - 2 subcategories
- Preferences: Strong – Moderate – Slight
- Multiple choice test

http://www.engr.ncsu.edu/learningstyles/ilsweb

They like innovation
They prefer verbally presented
Let’s think it through first
They absorb material almost randomly

- Sensing
- Intuitive
- Visual
- Verbal
- Active
- Reflective
- Sequential
- Global

They like innovation
They prefer verbally presented
Let’s think it through first
They absorb material almost randomly
Phase 2: Problem Detection and Solution

- **Main Goal:**
  - Analyzing the data to detect problems
  - Defining strategies to minimize the problems
Phase 2: Problem Detection and Solution

• Determining factors
  - Time overlap
    (high, intermediate, low)
  - Language difference
    (high, high-intermediate, intermediate, low-intermediate, low)
  - Cultural difference
    (high, intermediate, low)
  - Group Cognitive Style
    (type1, type2, type3)
Phase 2: Evaluating time overlap (Task 1)

Linguistic tags are easy to remember and to refer to.
Can be reused in different projects.
## Phase 2: Evaluating Language Difference (Task 1)

### Form 7: Degree of knowledge of a common language

<table>
<thead>
<tr>
<th>Possible common language</th>
<th>Language: .....................................................................................................................</th>
<th>Possible answer(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All the stakeholders are from the same country</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td>Stakeholders don’t share the mother language but they have a high level of knowledge about the chosen common language.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>High-Intermediate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stakeholders share the mother language but they are from different countries.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stakeholders don’t share the mother language but they have a high-intermediate level of knowledge about the chosen common language.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intermediate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stakeholders don’t share the mother language but they have at least an intermediate level of knowledge about the chosen common language.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low-Intermediate</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stakeholders don’t share the mother language but they have at least a low-intermediate level of knowledge about the chosen common language.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Low</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Stakeholders don’t share the mother language and all of them have a low level of knowledge about the chosen common language.</td>
<td></td>
</tr>
</tbody>
</table>
**Phase 2: Analyzing Cultural Difference (Task 1)**

**Hofstede model: Five dimensions**
- Power Distance Index (PDI)
- Individualism (IDV)
- Uncertainty Avoidance Index (UAI)
- Masculinity (MAS)
- Long-Term Orientation (LTO)

<table>
<thead>
<tr>
<th>Country</th>
<th>PDI</th>
<th>IDV</th>
<th>UAI</th>
<th>MAS</th>
<th>LTO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>49</td>
<td>46</td>
<td>56</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>36</td>
<td>90</td>
<td>61</td>
<td>51</td>
<td>31</td>
</tr>
<tr>
<td>Austria</td>
<td>11</td>
<td>55</td>
<td>79</td>
<td>70</td>
<td></td>
</tr>
<tr>
<td>Belgium</td>
<td>65</td>
<td>75</td>
<td>54</td>
<td>94</td>
<td></td>
</tr>
<tr>
<td>Brazil</td>
<td>69</td>
<td>38</td>
<td>49</td>
<td>76</td>
<td>65</td>
</tr>
<tr>
<td>Canada</td>
<td>39</td>
<td>80</td>
<td>52</td>
<td>48</td>
<td>23</td>
</tr>
<tr>
<td>Chile</td>
<td>63</td>
<td>23</td>
<td>28</td>
<td>86</td>
<td></td>
</tr>
<tr>
<td>China</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>118</td>
</tr>
<tr>
<td>Spain</td>
<td>57</td>
<td>51</td>
<td>42</td>
<td>86</td>
<td></td>
</tr>
</tbody>
</table>
• Group types

  ▪ **Type 1**
    No strong preferences in the team.

  ▪ **Type 2**
    There are strong preferences but not at the opposite ends of the same category.

  ▪ **Type 3**
    There are strong preferences at the opposite ends of the same category (conflict of preferences)
Phase 2: Defining strategies according to measured factors (Task 2)

(A) Cultural difference?
- Low
- Medium to High
  - Learning about Culture

(B) Same country?
- NO
  - Use of Ontologies
- YES
  - Intermediate to High
    - Common Language?
    - Low
      - Recommend asynchronous technology (C4)
    - High
      - Group type?
        - 1
          - Strategy C1
        - 2
          - Strategy C2
        - 3
          - Strategy C3

(C) Time overlap?
- Low
  - Recommend asynchronous technology (C4)
- Medium to High
- (A) Training to minimize cultural differences
- (B) Use of Ontologies
- (C) Technology Selection

<table>
<thead>
<tr>
<th></th>
<th>Inadequate communication</th>
<th>Time separation</th>
<th>Cultural diversity</th>
<th>Knowledge management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimizing cultural differences</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>Use of Ontologies</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Technology selection</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>
(B) Use of Ontologies

• Ontologies’ importance during requirements elicitation processes
  ▪ Clarify the structure of knowledge
  ▪ Reduce conceptual and terminological ambiguities

• Using a domain ontology as part of the elicitation process helps stakeholders’ communication
  ▪ Sets the basic vocabulary
  ▪ Clarifies meaning for each new word
Empirical Validation

- 16 students from UCLM (Spain)
- 8 software engineers/researchers from UNCOMA (Argentina)
- Randomly distributed

<table>
<thead>
<tr>
<th>Gw_+</th>
<th>O_+</th>
<th>O_-</th>
</tr>
</thead>
<tbody>
<tr>
<td>G4 G8</td>
<td>G6 G3</td>
<td></td>
</tr>
<tr>
<td>G1 G7</td>
<td>G2 G5</td>
<td></td>
</tr>
</tbody>
</table>
Empirical Validation

- 8 Teams = 2 analysts (Spain) + 1 client (Argentina)
- Problem: to obtain the requirements from the client
- 1 week to interact and write the software requirement specification

- We collected
  - Copies of emails and records of audio and chat conversations
  - Post-experiment questionnaire
• “Do you think the ontology was useful in improving communication in your team?”

Empirical Validation

0 useless
1 slightly
2 indiffer.
3. Useful
4. very
Empirical Validation

The bar chart shows the distribution of answers for different roles: 'Users' and 'Analysts'. The x-axis represents the roles, and the y-axis represents the number of answers. The chart includes four categories: 0, 1, 2, and 4.
Empirical Validation

• “How good do you think the quality of communication in your team was?”

Group 0 - not ontology
Group 1 - with ontology

0 very bad
1 bad
2 acceptable
3 good
4 very good
• Quality of the **software requirements specification** from the point of view of external reviewers

![Box Plot](Spreadsheet_Statistica_valoracion_expertos.xls)  
Group 0-not ontology  
Group 1- with ontology
Conclusion of the Validation

Ontologies used as a communication facilitator during a global requirement elicitation process

- Most people considered the ontology to be useful and very useful (especially analysts)
- However, with regard stakeholders’ satisfaction and productivity the results did not coincide with our expectations
- The language difference between the stakeholder are not very great
$H_{0,2}$: Stakeholders’ satisfaction about communication ($Gw_+, Gw_-$)

- $H_{0,2}$: Using groupware tools according to the cognitive profile of stakeholders does not affect stakeholders’ satisfaction with communication during the requirements elicitation process.
- $H_{1,2} = \neg H_{0,2}$

<table>
<thead>
<tr>
<th></th>
<th>comun21A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mann-Whitney U</td>
<td>28,000</td>
</tr>
<tr>
<td>Wilcoxon W</td>
<td>106,000</td>
</tr>
<tr>
<td>Z</td>
<td>-2.860</td>
</tr>
<tr>
<td>Asymp. Sig. (2 tailed)</td>
<td>.004</td>
</tr>
<tr>
<td>Exact Sig. [2*(1-tailed Sig.)]</td>
<td>.010(a)</td>
</tr>
</tbody>
</table>

a Not corrected for ties.
b Grouping variable: groupware

**We reject the null hypothesis $H_{0,2}$ (99%)**
Conclusions

• Global software development is affected by many factors which complicate communication

• We propose a framework which focuses on predicting problems and proposing strategies to avoid those problems

• We are currently evaluating the different strategies
Analyzing Ontology as a Facilitator During Global Requirement Elicitation

Gabriela N. Aranda

Aurora Vizcaíno

Mario Piattini

+GIISCo Research Group
Universidad Nacional del Comahue
Argentina

*Alarcos Research Group
Universidad de Castilla-La Mancha
Spain
Phase 2: Problem Detection and Solution

Groupware tools
Preferente rules

(MRe, MSe, VVe, SGl)

Strongly Verbal

(MAc, SSe, SVe, MSq)

Strongly Visual

(MAc, VSe, VVi, VSq)

1-Email
2-Chat
3-Video

1-Video
2-Chat
3-Email

CHAT
A Groupware tools Selection methodology

1. Obtaining people’s styles by F-S test
   - Obtaining people’s preferences about Groupware Tools
     - Applying a machine learning algorithm to examples to obtain preference rules

2. Applying F-S test to virtual teams members
   - Automatic analysis and selection of technology for the given virtual team
(A) Training to minimize cultural differences

- Cultural differences cannot be avoided
  - but stakeholders can learn about the differences of the other culture.

- Being trained about cultural diversity is crucial
  - To be aware of normal behaviour in other cultures
  - To be conscious of our own behaviour (what can be offensive or misunderstood)

- Strategy: Virtual mentoring (simulation)
  - virtual actors become an interesting way for motivating stakeholders in foreign language training and cultural familiarization (Sims, 2007)
• Variety of technology
  ▪ Groupware Tools
    • e-mail, chat, instant messaging, videoconference, audio conference, shared whiteboards, …
  ▪ Requirements Elicitation Techniques
    • Interviews, software prototypes, use cases and scenarios, observation, …

• Selection based on stakeholders’ cognitive characteristics
  ▪ Machine learning algorithms applied on preference examples, using F-S Learning Style Model
Cognitive Style

- Strongly visual
- Moderately active
- Slightly sensing
- Slightly global

- Moderately verbal
- Slightly reflexive
- Moderately intuitive
- Moderately sequential

Preference Rules

Shared whiteboards (+++)
videoconference (+++)

Interviews (+)
Prototype (++)

Technology Selection
Empirical Validation

• **Independent Variables**
  - Ontology (Yes/No)
  - Groupware (Most suitable/Less suitable)

• **Fixed Variables**
  - Cultural difference (low)
  - Language difference (low-intermediate)
  - Time separation (4 hours = intermediate)
  - Type 2 Groups (strong preferences without conflict)

• **Dependent Variables**
  - Stakeholders’ satisfaction with communication during the process
  - Stakeholders’ satisfaction with the product (SRS)
  - Software Requirements Specification (SRS) quality
**H_{0,1}: Using a domain ontology does not affect stakeholders’ satisfaction with communication during the requirements elicitation process.**

- \( H_{1,1} = \neg H_{0,1} \)

\[ \begin{array}{|c|c|}
\hline
\text{comun21A} & \\
\hline
\text{Mann-Whitney U} & 66.500 \\
\text{Wilcoxon W} & 144.500 \\
\text{Z} & -.357 \\
\text{Asymp. Sig. (2 tailed)} & .721 \\
\text{Exact Sig. [2*(1-tailed Sig.)]} & .755(a) \\
\hline
\end{array} \]

- a Not corrected for ties.
- b Grouping variable: ontology

*We cannot reject the null hypothesis*
H_{0,2}: Using groupware tools according to the cognitive profile of stakeholders does not affect stakeholders’ satisfaction with communication during the requirements elicitation process.

\[ H_{1,2} = \neg H_{0,2} \]

We reject the null hypothesis \( H_{0,2} \) (99%)
H_{0,3}: There is no interaction effect between using a domain ontology and groupware tools according to the cognitive profile of stakeholders, concerning stakeholders’ satisfaction with communication during the requirements elicitation process.

H_{1,3} = \neg H_{0,3}

**Table:**

<table>
<thead>
<tr>
<th></th>
<th>comun21A</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chi-cuadrado</td>
<td>8,433</td>
</tr>
<tr>
<td>gl</td>
<td>3</td>
</tr>
<tr>
<td>Sig. asintót.</td>
<td>0,038</td>
</tr>
</tbody>
</table>

a) Kruskal-Wallis Test  

b) Variable de agrupación: treatment

**We reject the null hypothesis H_{0,3} (95%)**