



**UNIVERSITAT POLITÈCNICA  
DE CATALUNYA**  
**BARCELONATECH**

Facultat d'Informàtica de Barcelona

# **Crowd outsourcing for software localization**

**A Collaboration with CA Technologies.**

**Supervisor:** Victor Muntés

**Tutor /Professor:** Josep Larriba-Pey

**Student:** Andrea Gritti

**Master:** Master in Information Technology

**University :** Universitat Politècnica de Catalunya

**BarcelonaTech**

09/01/2012

# Crowd sourcing for Translation and Software Localization

Master Thesis in order to acquire the Master in Information Technology

by

Andrea Gritti

Academic year 2011 - 2012

## **Abstract**

This work studies the capability of Crowd sourcing related to translation and software localization and the quality obtained by the use of a crowd sourcing methodology. This work is performed within the collaboration of CA Labs, Europe, and is specifically focused on the design of a crowd sourcing platform able to guarantee high quality in translation, and to comply with industrial aspects of translation.

Moreover, a prototype of the designed platform has been developed and has been used to run some experiment with a reduced and controlled crowd, to test the potentiality of translation done by a not homogeneous group of users.

The reasons, challenges, road-map and results obtained in this work are described in detail in this document.

## **Keywords**

Crowd sourcing, distributed computing, software localization and translation,

## Content

<b>1</b>	<b>INTRO.....</b>	<b>3</b>
<b>2</b>	<b>PROJECT BACKGROUND.....</b>	<b>6</b>
<b>2.1</b>	<b>Translation and Software Localization.....</b>	<b>8</b>
<b>2.2</b>	<b>CA Technologies.....</b>	<b>11</b>
<b>3</b>	<b>CROWD SOURCING.....</b>	<b>14</b>
<b>3.1</b>	<b>Crowd sourcing, what is all about?.....</b>	<b>15</b>
<b>3.2</b>	<b>Crowd sourcing studies.....</b>	<b>20</b>
<b>3.3</b>	<b>Crowd sourcing for Translation and language studies.....</b>	<b>23</b>
<b>3.4</b>	<b>Existing systems for Crowd sourcing.....</b>	<b>25</b>
3.4.1	Amazon mechanical turk.....	25
3.4.2	Duolinguo.....	29
3.4.3	Icanlocalize.com.....	31
3.4.4	Crowdfower.....	32
3.4.5	InboxDollar.....	34
<b>4</b>	<b>DESIGN.....</b>	<b>36</b>
<b>4.1</b>	<b>Use case.....</b>	<b>38</b>
<b>4.2</b>	<b>Logical View.....</b>	<b>44</b>
<b>4.3</b>	<b>First Version.....</b>	<b>45</b>
4.3.1	Description of the system.....	45
<b>4.4</b>	<b>Final Version.....</b>	<b>53</b>
4.4.1	Description of the system.....	55
<b>4.5</b>	<b>Ranking System.....</b>	<b>66</b>
<b>4.6</b>	<b>Action-Verification Unit.....</b>	<b>67</b>
<b>5</b>	<b>PROTOTYPING.....</b>	<b>69</b>
<b>5.1</b>	<b>Technology.....</b>	<b>72</b>
<b>5.2</b>	<b>Overview of the application's web flow.....</b>	<b>74</b>
<b>5.3</b>	<b>Data Base.....</b>	<b>79</b>
<b>5.4</b>	<b>Parser.....</b>	<b>83</b>
<b>5.5</b>	<b>Post Edit phase.....</b>	<b>86</b>
<b>5.6</b>	<b>Verification of Post edit phase.....</b>	<b>90</b>

5.7	Post Edit phase 2.....	95
5.8	Review.....	99
5.9	Verification of Review.....	102
5.10	Review phase 2.....	106
<b>6</b>	<b>RESULT ANALYSIS.....</b>	<b>108</b>
6.1	Result Analysis.....	111
6.1.1	Results.....	113
6.2	Survey.....	114
<b>7</b>	<b>COMPARISON.....</b>	<b>116</b>
7.1	Discussion.....	118
7.2	Crowd Sourcing Adoption.....	124
7.2.1	Cost Calculation.....	126
7.2.2	Organizational Change.....	127
<b>8</b>	<b>CONCLUSION.....</b>	<b>128</b>
8.1	Advices and thoughts.....	130
8.2	Ranking module suggestion.....	133
8.3	Knowledge acquired and Reflections.....	135
8.4	Future works.....	136
<b>9</b>	<b>GLOSSARY.....</b>	<b>137</b>
<b>10</b>	<b>BIBLIOGRAPHY.....</b>	<b>138</b>

# 1 Intro

Internet is being used by a steadily increasing number of users, connected from every corner of the world, creating an incredibly various mix of different knowledge, culture and languages which blend together on the same global net.

This variety of human culture and expectations, pose a barrier to software creations and distribution, since many people around the world are not able to understand any foreign language.

At the same time, information systems are also on a unstoppable evolution and continual improvement. Technology is on a constant challenge against problems that, nowadays, are still computationally difficult to be addressed by computers, while they are relatively easy for human intelligence to be solved instead. One of those open problems is translation and natural language processing: even the most recent machine translation technology can not produce, most of the times, results comparable, for quality and style, to a text translated by a human being.

Crowd sourcing is a new form of collaborative work between computer systems and people that has been arising during the last few years. Crowd sourcing aims at generating a scalable and flexible collaboration methodology between computers and humans beings, where users are in charge to solve small and easy tasks, while computers are in charge of orchestrating, collecting and composing all the results in order to solve a big problem.

Crowd sourcing is nowadays under the magnifying glass, and its capabilities are far from being fully understood; This magnifying glass has been used in this project to study crowd sourcing's potentiality, especially related to Translation and Software Localization process. Consequently, I have learned the current Translation and Localization process, and a crowd sourcing platform suitable for this job has been designed. Moreover, during the project some evidence to support the hypothesis that crowd sourcing may be a possible solution to many

of the current limitations in the Translation process.

This document starts with the description of the localization and translation process in the software industry. To help the reader to contextualize the development of the project, I describe the current limits and difficulties in translation and localization work, using a concrete example of the Translation and Localization department of CA Technologies. At any rate, the problems found during the collaboration with CA Technologies can be easily shared to others companies.

Crowd sourcing technology is extensively described in chapter 3, explaining aspects and characteristics that may avail this methodology to become a valuable resource in addressing translation and localization problems.

After these theoretical introductions, chapter 4 presents the work done during the project, starting with the detailed description of the design process. One of the main objective of the project was the design of a crowd sourcing platform able to comply with some industrial aspects of translation.

After the description of the design, chapter 5 introduces the coding of a prototype of the system. A prototype has been developed, then used to run some experiments and analyze the capability of crowd translation.

Chapter 6 presents some aspects of the on running experiment. The experiment was expected to end by August 2012, but it is still running due to some delay. Anyway, some preliminary results are offered to the reader.

Chapter 7 makes profit of the knowledge acquired during this project to compare different crowd sourcing systems with the designed system. In addition, few arguments are presented for a crowd sourcing adoption strategy and procurement guideline.

Chapter 8 concludes the thesis by listing ideas, thoughts, pain and joy lived during those months.

This project has been supported by the collaboration between UPC and CA Labs Europe, which has proposed the topic, the objectives and it has been

providing guidance, technical knowledge and financial support for the completion of the project.

## 2 Project Background

This chapter wants to help the reader in contextualizing the project background by explaining some of the current challenges in the Translation and Localization world. Then CA Technologies is presented in chapter 2.2, within some of the problems and limitations found during the analysis.

Since this project is a collaboration between CA Labs Europe and UPC, a close relationship has been established among the Head of CA Labs Europe Victor Muntés, another student from UPC, me and Professor Josep Larriba, which is the tutor of the project.

The collaboration between academy and company has the main objective of studying the potentiality a crowd sourcing system can offer. CA Technologies is in search of tangible evidence that prove crowd sourcing may be a possible solution to deal with the current problems present in its Translation and Localization department. On the other hand, crowd sourcing is a new methodology which recently created interest in the academic world.

Efficient communication, smooth interaction among the team and project management was a crucial part in order to succeed. The project has a large scope, potentially reaching all the languages that can be found on Internet. Even during the project, many different languages were involved: Spanish and Catalan were the languages chosen for the experiment, while English has been the language used for daily communication among all the people involved in the project.

The project begun with a deep study of Crowd sourcing state of art, answering questions such as: what it has been used for, which results have been obtained and how were they obtained, which different architectures were used in similar experiments, trying to understand the motivations behind them.

After the prerequisites analysis and with the suggestion of CA Technologies translation department, a crowd sourcing system was designed, and the



architecture presented novel features such as:

- Automatic parallelization and distribution of tasks;
- Creation of a quality management layer above the crowd;
- A fair rewarding system for users;
- An adaptable and scalable sequence of steps in the translation process;

Summarizing, the project consisted in two main goals: the initial requirement analysis and design of a crowd sourcing platform, and the development of a prototype where to perform some experiments within a reduced crowd.

The project itself take profit of the courses attended at FIB, including but not limiting to: PESBD, PROP, ES1 and ES2.

## 2.1 Translation and Software Localization

Before being involved in this project, I used to consider Software localization as a minor aspect of software development. I used to think about localization just as the mere translation of manuals and whereabouts of a specific system or software. This was just lack of deep understanding in the existing issues behind software localization.

I have been lucky enough to realize what translation and software localization is all about, and I could understand that it is not simply as it may seem: Localization is the modification of a business, product, content, or such, so that it satisfies both the language and cultural differences of the targeted market.

Language and cultural differences are two faces of the same coin. Coherently, localization process is made of two main aspects:

1. Technical Localization = is the pure act of translating text in the target language. While this is not easy as it may seem, as there are plenty of examples where programs that operate well in one language, with one set of characters, crash in another. Non-English speaking users may receive unintelligible, awkward, or English-language-only help, tutorials, and documentation.

Although not always possible, yet some of these problems are, at least in principle, resolvable by some strategies suggested in [1] where the problems are addressed at root, by planning for internationalization at the earliest stage of software design and coding, by a consistent interaction of internationalizers, back-translator of translated texts and adequate beta testing of localized programs with native people.

2. Cultural localization = it is the process where text is adapted in a way that they seem fully consistent with the assumptions, values and outlooks to the target language.

In other words, a culturally localized program "should be

indistinguishable from a program written by members of that culture" avoiding situations where poor translation is the cause that lead users to commit errors.

Considering the persuasive use of technology in those days, and the different fields where it is used, the process of localization presents different problems for different kinds of programs and systems: while programs for scientific use may not present much problem of Cultural localization, when it comes to other systems such, in example, management support systems, educational programs and even OS, an incorrect localization process creates systems that are more prone to cultural misunderstanding. Whenever a program is approximately translated its usability may be reduced, resulting in a negative impression on the final user.

In other words, while the system may be of top quality, poorly translation would lower the quality perceived by customer, and so decreasing the overall value of a system and tamper the sales, when it comes to foreign countries.

Moreover, a misunderstanding created by an incorrect translation in the user interface or user guide of a product may cause user to unwittingly perform a wrong actions, potentially causing lost of information and affecting negatively a software product's performance and image.

But problems of poor localization are not strictly related to user experience and technical issues, and they may, sometimes, reach to touch political spheres.

As documented in [2] when Microsoft opened the first office in Beijing in 1992, it started introducing programs in Chinese-language that used the Mandarin character set used in pre-revolutionary China, and still used today in Taiwan (Taiwan had closer political interaction and more economical interchange with USA at that time).

Microsoft did this without knowing that the government of the People's Republic of China (PRC), after assuming power in 1949, had modified these traditional character set, introducing simplified Chinese characters, now universally used in China.

PRC authorities were therefore predictably offended by Microsoft's practice of translating software in Redmond, Washington, and committing such a lowbrow error. The PRC's attitude to Microsoft allegedly softened only after Microsoft president Bill Gates spent an extended period in China to develop a collaborative project to train local technicians at Chinese centers and universities, in order to develop simplified Chinese operating systems in China. As a result, Chinese Microsoft's employees were discovered to insert in some Microsoft programs some hidden anti-regime slogans, and this again strained Microsoft-PRC relations.

This situation may have led to worse consequences, if not a dramatic change in business was taken: The world's leading software company risked exclusion from the world's largest potential market, because of a "cultural" failure in understanding the current creed in modern China.

This example clearly describes how software localization is not just a basic requirement, often imposed by law, to export software systems to foreign countries, it is instead a complex and complicated process which plays a vital role in a globalized world.

## 2.2 CA Technologies

CA Technologies, is one of the largest software corporation in the world.

With a \$4.4 billion US\$ in revenue for fiscal year 2010, 150 offices in 45 different countries, CA Technologies provides work to more than 13,000 employees, all of them allocated onto a variegated range of software products that include anti viruses, mainframe applications, distributed computing, virtualization and cloud environments.

Because of its size and its presence in many different markets around the world, CA Technologies must create its software in a localized way, in order to satisfy the global demand to where it provides its products.

CA Technologies has been involved in software localization activities for more than 10 years, both from the user and the developer perspective, developing language translation technologies and tools to support translation processes, and it is currently investing proximately around 10M \$ per year to satisfy CA Technologies quality standards in translation.

CA Technologies has a localization team located in Cornellà de Llobregat, which is in charge of translating both software and user guides, and it is responsible of the localization process behind the commercialize of CA Technologies products in Europe. This team is composed of 20 professional translators and it is in charge of translating both User interfaces of software application and manuals for its products, from English to 4 different European languages such as: Spanish, French, Italian and German.

The current work methodology implies the use of the most recent Machine Translation technologies, and then an extensively post-edit of the MT output in order to reach the CA Technologies quality standards.

During the project, different meetings with employees from this department have been held. Within the collaboration of Patricia Palladini, head of the

translation and localization department, and through the analysis of the current situation and work done at this department, some problems were detected:

- Longer «time-to-market» periods for non-English versions of CA Technologies products: products translated to languages different from English are usually released three months after the English version, because the localization process is time-consuming.
- Changing workload management: the translation workload is not heterogeneous and there are some peaks during the year when a large number of products are released together. Hence, the localization teams cannot cope with the current methodology whenever those peaks show up, forcing to outsource part of the software localization.
- High cost of extending to market in countries where spoken languages are not translated by the localization department: in order to translate to a large number of languages, and open the possibilities for the company to explore new markets, the current approach does not work. Firstly, it is not easy to find translators for all minority languages. Secondly, it is economically expensive and thus unfeasible for the company to hire a team of translators for every language, especially for emerging markets in some countries where the number of products sold is not expected to be huge such it may be the case, for example, of middle east or east Europe.
- High cost of software localization for languages that are currently translated: CA Technologies invests several million dollars in localization per year both in internal and outsourced localization. This causes several products not to be considered for internationalization even for common languages such as Spanish, German or French.

In other words, the current situation presented problems and limitations: high cost of translation, longer time to market for non English products and difficulty to translate do less popular languages.

In consequence, some objectives and scopes have been defined accordingly with CA Technologies, in order to address and solve the over mentioned problems:

1. **Reducing the cost of translation:** part of the translation is made by CA Technologies internal team and another important part is outsourced to language vendors: approximately, 23% of this budget is spent in linguists and another 23% in translation services outsourcing. Outsourcing represent an important expense for CA technologies, which it should be reduced, or even avoided;
2. **Reducing the delay to market of the company's products:** CA technologies wants to achieve simultaneous delivery for English and not English products, and even the current Outsourcing partner are not able to provide this;
3. **Obtaining a translation output acceptable under quality standards:** CA technologies wants its quality standards to be reached, and, thus, quality management is a central point in the development of the project;
4. **Increasing the capacity to translate to other languages:** CA Technologies is currently translating software from English to around 10 different languages. However, being able to translate to a larger number of languages could open the possibilities for the company to explore new markets.

All those aspects have been taken in the highest consideration during the development of this project, and the problems found have been addressed one by one during the design phase of the system.

### **3 Crowd Sourcing**

*What exactly is crowd sourcing?*

This chapter begins with chapter 3.1 which introduce crowd sourcing by mentioning its born, evolution and diffusion. Then chapter 3.2 describes the current state of art in crowd sourcing studies, while chapter 3.3 talks about crowd sourcing systems and its use for language studies. The chapter ends with a panoramic on the most important crowd sourcing systems nowadays.



### 3.1 Crowd sourcing, what is all about?

There is not a standard definition of crowd sourcing, and many different terms refer to it: Crowd computing, Crowd sourcing, Collaborative intelligence, distributed thinking, Crowd casting and so on.

While describing some slightly different aspects and characteristics, all terms are under the same umbrella that covers the methodology where computation is performed by humans within the interaction of computers.

The authorship of term "Crowd sourcing" is given to Jeff Howe, who firstly coined it in his article: «The rise of crowd sourcing»<sup>\*</sup> published on Wire magazine in June 2006. In this article Howe describes new internet trends where people were starting to use the global net in order to solve problems that already existed and traditionally solved in others manners. The article specifically talks about the gathering of good quality photos, and compare the traditional situation where, whenever companies or individuals needed good quality pictures, the only possible way was to contract professional photographers or deal with photography agencies, while the advent of istockphoto.org, unprofessional but skilled photographer started to share their good quality pictures on-line, creating a market of good quality photos, worldwide accessible.

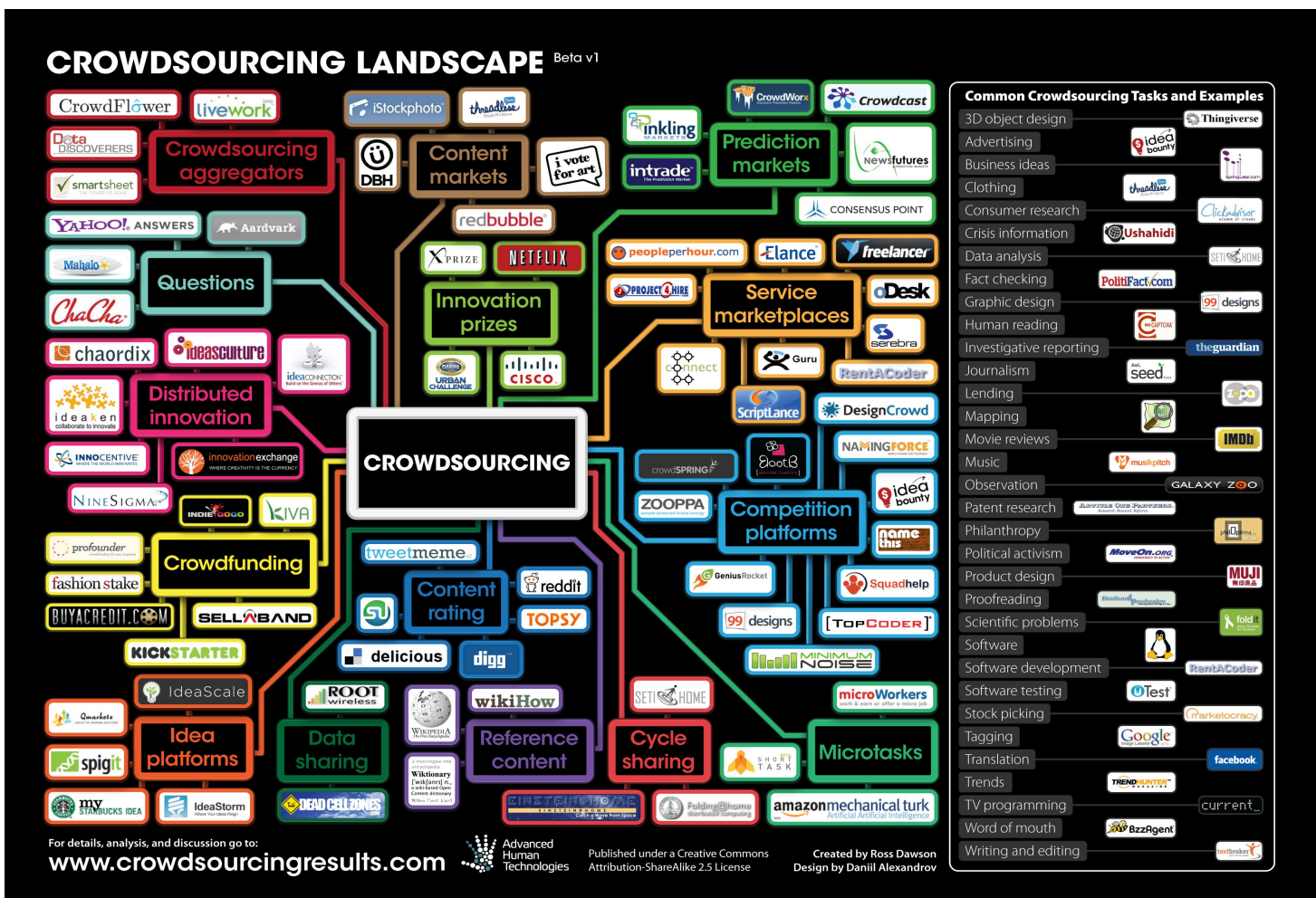
In this visionary article, Howe is bright enough to describe the collateral effects that crowd sourcing was causing. While the traditional outsourcing of jobs to other countries has been often accused to relocate and even steal jobs from local workers, crowd sourcing moves this accusation from a principle of locality to a principle of skill sets. In other words, the core issue was no longer the fact that producing in another part of the world was cheaper, but the fact that a larger market was accessible, wherefore an increased offer made quality cheaper due to a larger number of competitors.

<sup>\*</sup>[\[http://www.wired.com/wired/archive/14.06/crowds.html\]](http://www.wired.com/wired/archive/14.06/crowds.html)

Howe coined the term Crowd sourcing by mixing an existing term such "outsourcing", and adding the term Crowd, where with crowd he referred to an undefined and usually large network of users.

So, crowd sourcing can be considered as an outsourcing method where task that traditionally were performed inside a company with specific roles, they are divided and outsourced to a crowd of users which, usually, answer to an open call request for completing tasks in exchange of a reward. Users complete tasks using different kinds of coadjuvant tools and usually perform simple tasks for a lower price.

Since 2006, crowd sourcing has become more popular, and it has evolved in many different services, as it can be seen from the following picture.



From the image, it is easy to see how the term crowd sourcing includes many systems, even quite different between each others; many studies have tried to categorize this foggy cloud.

In [3] authors provide a survey on the literature of crowd sourcing, and try to categorize crowd sourcing systems in 4 different categories, according for what crowd sourcing is used for:

Application: are the crowd sourcing systems where usually a requester creates a task and propose it to the crowd as an open call proposal. Tasks can be of any kind.

Algorithms: those systems use crowd sourcing to compute some information, and those information are processed within a defined algorithm.

Performances: are the systems that use the crowd to study user's behavior under different criteria. Rather as social experiment and investigation, more than technology research.

Data sets: Crowd sourcing is an efficient, fast and cheap method to obtain large volume of human produced data. Systems that use the crowd for this data production are categorized in this area.

Anyhow, at present, there is not any general-accepted crowd sourcing categorization. Different platform which embeds task-based services to support crowd sourced activities are labeled as crowd sourcing systems, while they may not exactly be considered as crowd sourcing systems.

This unclear definition made approaches to this technology to be naive. In the paper [4] authors describe in detail some main phases of a crowd sourcing process, and provide a series of minimum requirements and guideline to correctly approach crowd sourcing services.

To conclude with, even if there is not a globally accepted and standard definition for crowd sourcing, I consider the most complete definition the one given in [5] where authors have been collecting different definitions of crowd sourcing, analyzing what they had in common, and creating an unique one. The final

definition is the following:

*“Crowd sourcing is a type of participative online activity in which an individual, an institution, a non-profit organization, or company proposes to a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task. The undertaking of the task, of variable complexity and modularity, and in which the crowd should participate bringing their work, money, knowledge and/or experience, always entails mutual benefit. The user will receive the satisfaction of a given type of need, be it economic, social recognition, self-esteem, or the development of individual skills, while the crowd sourcer will obtain and utilize to their advantage that what the user has brought to the venture, whose form will depend on the type of activity undertaken.”*

The characteristics of the definition are:

- a) There is a clearly defined crowd;
- b) There exist a task with a clear goal;
- c) The recompense received by the crowd is defined;
- d) The requested is clearly identified;
- e) The result that crowd sourcer receives from the crowd is clear;
- f) It is an online assigned process of participative type;
- g) It uses an open call of variable extent;
- h) It uses Internet;

According to this definition, it can be noticed as some of the systems reported in the previous image (as in example Wikipedia) can not be considered crowd sourcing systems.

**Table 4. Verification of the definition. Source: author**

	<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>	<b>e</b>	<b>f</b>	<b>g</b>	<b>h</b>
Wikipedia	+	+	+	-	-	+	-	+
InnoCentive	+	+	+	+	+	+	+	+
Threadless	+	+	+	+	+	+	+	+
Amazon Mechanical Turk	+	+	+	+	+	+	+	+
ModCloth	+	+	+	+	+	+	+	+
YouTube	+	-	-	-	-	-	-	+
Lánzanos	+	+	+	+	+	+	+	+
Delicious	+	-	-	-	-	-	-	+
Fiat Mio	+	+	+	+	+	+	+	+
iStockPhoto	+	+	+	+	+	+	+	+
Flickr	+	-	-	+	-	-	-	+

### 3.2 Crowd sourcing studies

Compared to greater literature on similar technology, as it may be distributed computation and distributed systems, studies on crowd sourcing recently began.

Some systems such Amazon Mechanical Turk\* are now really popular and studied extensively. Aforesaid kind of systems are designed to give businesses and developers access to an on-demand, scalable workforce by creating tasks to be solved by any user in the Internet (Amazon Mturk is better described in chapter 3.4.1) .

Being one of the first to be developed allowed Mturk to become one of the most populated system, and because of its intrinsic features[6], Mturk has been used in many different studies which involve crowd sourcing.

One of the most studied aspects of crowd sourcing is quality of results. Since it seems to be totally aleatory, up to now there are not any well-known criteria which influence the quality of the result obtainable from crowd sourcing. In Mark D. Smucker[7] authors did some experiment on a task of Information Retrieval using crowd sourcing, and in their study is shown how the relevance judging behavior of a crowd-sourced group of participant is more inaccurate compared to behavior of laboratory participant.

At the same time, it seems that quality obtainable from crowd sourcing strongly changes accordingly to the nature of task which is crowd sourced; in example, Tingxin Yan, Vikas Kumar, Deepak Ganesan [8] proposes a system called "CrowdSearch", which is an image search method for mobile phones, where automated image search is combined with real-time human validation of the search results. In this case, crowd sourcing workers were able to achieve over 95% precision in judging images across multiple categories, making crowd sourcing to be fast, cost effective, really efficient and well fitted for this task.

\*[[www.Mturk.com](http://www.Mturk.com)]

Consequently, it can be seen how some aspects of a specific task make the task itself more or less suitable to be solved with crowd sourcing. This point will be addressed in a more exhaustive manner in chapter 4.6.

During the initial period of the project, the study of crowd sourcing state of art made clear many aspects that are tampering with the diffusion of crowd sourcing; one of those is Law. Law is not keeping up with the evolution of crowd sourcing, which is meant to change the traditional methodology of outsource and, nowadays, Law presents many unaddressed issues such employment law, patent inventorship, data security and ownership, copyright and so on, aspects which are not well regulated by law in the case of crowd sourcing, as it is well explained in [9].

Another important aspect which have been studied is motivation. What are the main reasons that lead users to participate in crowd sourcing systems? The answer to this question is the topic of many studies, such, in example in [10] is described as users located in in underdevelopment countries uses crowd sourcing systems as primary income resource, fulfilling basic needs of many people, especially in India. In Dana at al [11] motivation of crowd sourcing workers have been approached with the objective of discover how motivation influences results quality, and in which way. The experiment done by Dana at al is particularly representative of crowd motivation, especially because it touches an uncovered nerve: the ideological support for science for common well-being. The experiment asked to crowd users to recognize certain patterns in a group of photos. Users that accept to participate were divided in two different groups which received different introductory instructions and information about the tasks: the first group was informed that user responsibility was to help the tumoral research by finding possible tumoral cells; the second group was simply informed on what and how they had to do, without explaining the context, neither the reason of their work. Moreover, the amount of money paid for task was decreasing for each task completed, in order to see if the simply philanthropic motivation was enough to motivate users.

As it can be thought intuitively, the two groups provided significantly different results: more accurate and with a general higher quality came from the first group, while sketchy and inaccurate results were delivered from the second group. Ergo, motivations to accomplish a certain task seems to be a fundamental issue that lead to different quality of produced work.

In spite of intuitive aspects, there are other factors that influence the result's quality which are not intuitive at all: Downs et al. [12] screened crowd workers motivation by asking Amazon Mturk users to answer some general questions, which authors already knew the answer. Experimental results showed that students and unemployed users seem to be more likely to take the task seriously and provide more thoughtful answers, compared to other workers. Besides, men over 30 and women of any age were much more likely to qualify for a better answer.

Summarizing, researchers have studied crowd sourcing systems from a less technical view, focusing their effort on sociological aspects and behavioral patterns that could lead to a better quality obtainable from crowd.



### **3.3 Crowd sourcing for Translation and language studies**

Crowd sourcing seems to be very suitable for language studies, especially because of its own nature of multilingual and multicultural environment. In fact, many experiments described in recent papers have used crowd sourcing techniques to evaluate machine translation(MT) outputs, while in other studies researchers have investigated some mechanisms for human translation which may be suitable for crowd sourcing.

Starting from year 2009, in [13] authors describe an innovative approach in evaluation of Machine Translation outputs. They replaced the traditional BLUE metrics [14] with human evaluation obtaining evaluation of MT quality through reading comprehension performed by a large number of users on Amazon mechanical turk. This evaluation provides a better evaluation of the capacity of Machine Translation, but it was rarely done before the advent of Mturk, due to the expensive and infeasibility of gathering, sending, receiving and coordinating a vast number of users.

As said before, crowd sourcing is not used just for evaluation of Machine Translation, but also to actually perform translation. For example, Omar et al [15] suggest a collection of criteria which are proven to influence the quality of a translation obtained with crowd sourcing. A set of translation were collected and classified according to criteria such, in example, worker's country, native language and culture, LM perplexity of the translation, edit rate from other translations. Then authors gave a different weight to each criteria, they ordered the set of translation accordingly to those criteria, and compared the translation with some golden translation provided by professional translators. The result were very encouraging, especially for the translation ranked at the top of the list, where quality provided by native speakers was very close to quality provided by professional translators.

Nevertheless, the experiment was, in a certain sense, limited under some circumstances such, in example, the length of the translated text. That is to say

the overall objective of this experiment was to discover the existence of quality in translation made in crowd sourcing systems, rather than finding which criteria influence it, neither the study of methodologies able to guarantee quality in crowd sourcing translation.

In [16] and [17] author's attention focuses on the interaction between Machine Translation and Crowd sourcing systems to see how this interaction may lead to an improvement of Machine Translation technology. Those studies highlight which part in natural language processing done by Machine Translation was more influenced by the use of crowd, and in which way.

For example, the interaction between MT and Crowd sourcing has been used for word alignment [18], a vital task for MT performance, by incorporating manual word alignments done with Amazon Mechanical Turk (Mturk).

In Qin Gao[19] and Negri et al [20] Amazon mechanical turk is used for creating text corpora used to enrich Statistical Machine Translations.

One of the most interesting tool that exploits crowd sourcing potentiality in a really neat manner is presented in [21] where Bernstein et al describe a novel pattern called «Find-Fix-Verify» composed by different stages where independent agreement and collective voting are used to elect the best translation among a set of possible translations.

Many papers expose the problem that crowd sourcing quality should not be delegated to crowd users capacity and motivation, and it actually exists the clear need of some automatic quality assurance method, which the actual lack of it is impeding crowd sourcing system to become compliant with industrial aspects, especially when it comes to translation and software localization.

### 3.4 Existing systems for Crowd sourcing

This chapter introduces the most important crowd sourcing systems, and provides an overview of the actual panorama for translation and software localization.

#### 3.4.1 Amazon mechanical turk

Amazon Mechanical Turk is one of the first crowd sourcing internet marketplaces, launched by Amazon in November 2005.

Originally, Amazon Mechanical Turk was not designed for any particular nor specific task (as it may be translation) and it was meant to give businesses and developers access to an on-demand, scalable workforce.

Amazon Mechanical Turk allows registered users to publish tasks of any kind, and offer rewards in exchange for a completed task; those tasks are called Human Intelligence Tasks (HIT) and are usually simple, quick and repetitive.

HITs are published on the Mturk website as an open call, where any other registered user can accept a task and accomplish it in order to receive the rewards. Users in Amazon Mechanical Turk are named Turkers.

#### Make Money by working on HITs

HITs - *Human Intelligence Tasks* - are individual tasks that you work on. [Find HITs now.](#)

**As a Mechanical Turk Worker you:**

- Can work from home
- Choose your own work hours
- Get paid for doing good work



#### Get Results from Mechanical Turk Workers

Ask workers to complete HITs - *Human Intelligence Tasks* - and get results using Mechanical Turk. [Get started.](#)

**As a Mechanical Turk Requester you:**

- Have access to a global, on-demand, 24 x 7 workforce
- Get thousands of HITs completed in minutes
- Pay only when you're satisfied with the results



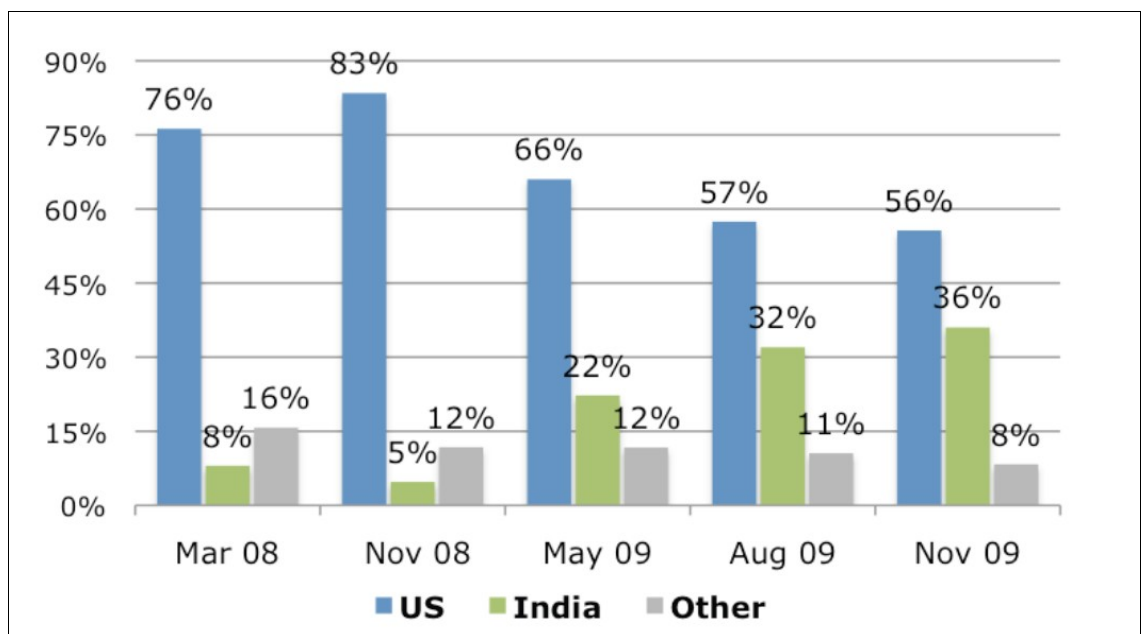
Amazon Mechanical Turk offers tasks which are usually trivial to resolve, rapid to be performed and paid very poorly and with no variation depending on the

quality of the results, and, if the Requester is not satisfied with the quality provided by the Turker, he can reject the task and do not pay the the Turker for the work done and time spent on the task.

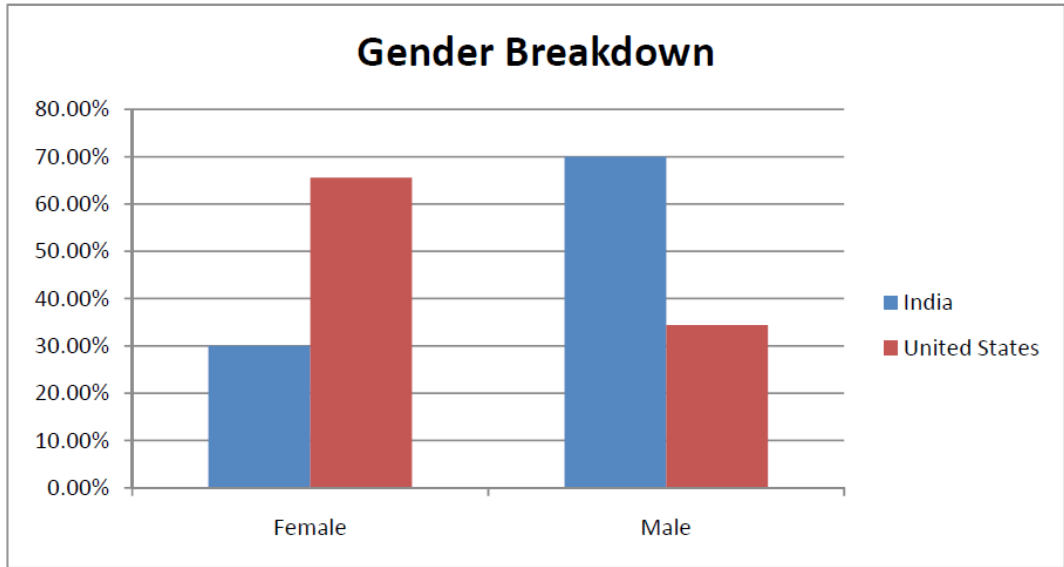
Tasks are paid from 0.01 \$ for a five second task, up to few dollars for more complicated tasks which can take minutes. Turkers earn on average 2.3 \$/hour in US and 1.7\$/hour in India. Even though Amazon Mechanical Turk offers some way to filter and organize tasks, such, in example, limit the access to HIT accordingly to nationality or country of Turker, Amazon Mechanical Turk is fundamentally a flat system where to offer and resolve easy peasy tasks.

Nevertheless, Amazon's popularity brought Amazon Mechanical Turk to become the most important crowd sourcing market; Amazon reports that the system has now more than 400,000 workers registered, and there are about 50,000-100,000 HITs to work on at any given time.

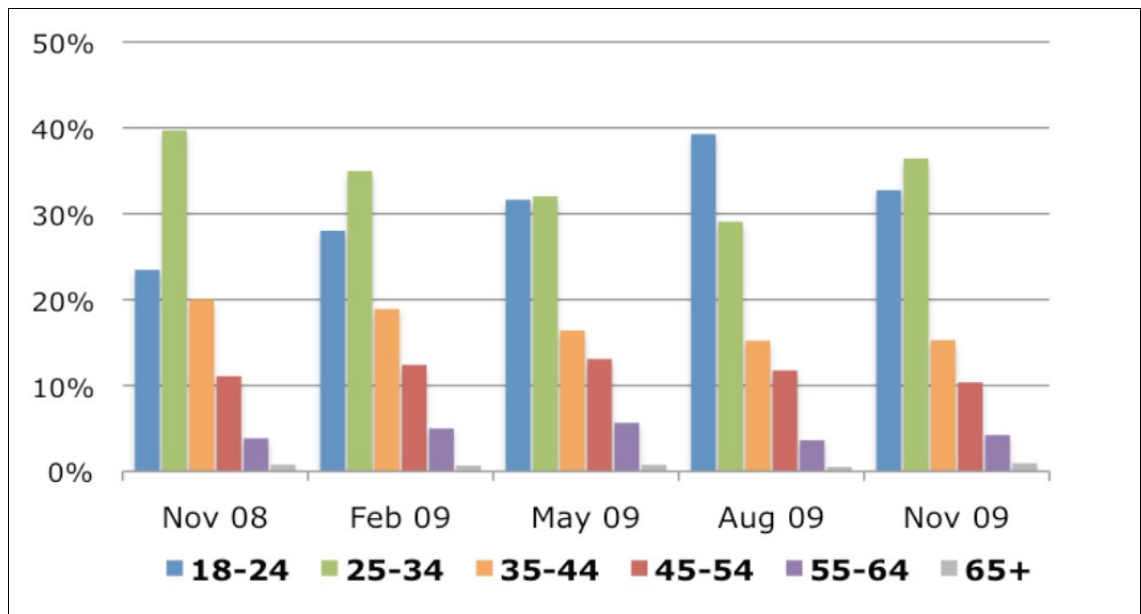
According to data reported in [22] and [23] Turker population has changed over time, shifting from a primarily moderate-income, U.S.-based workforce, towards an increasingly international group with a significant population of young, well-educated Indian workers.



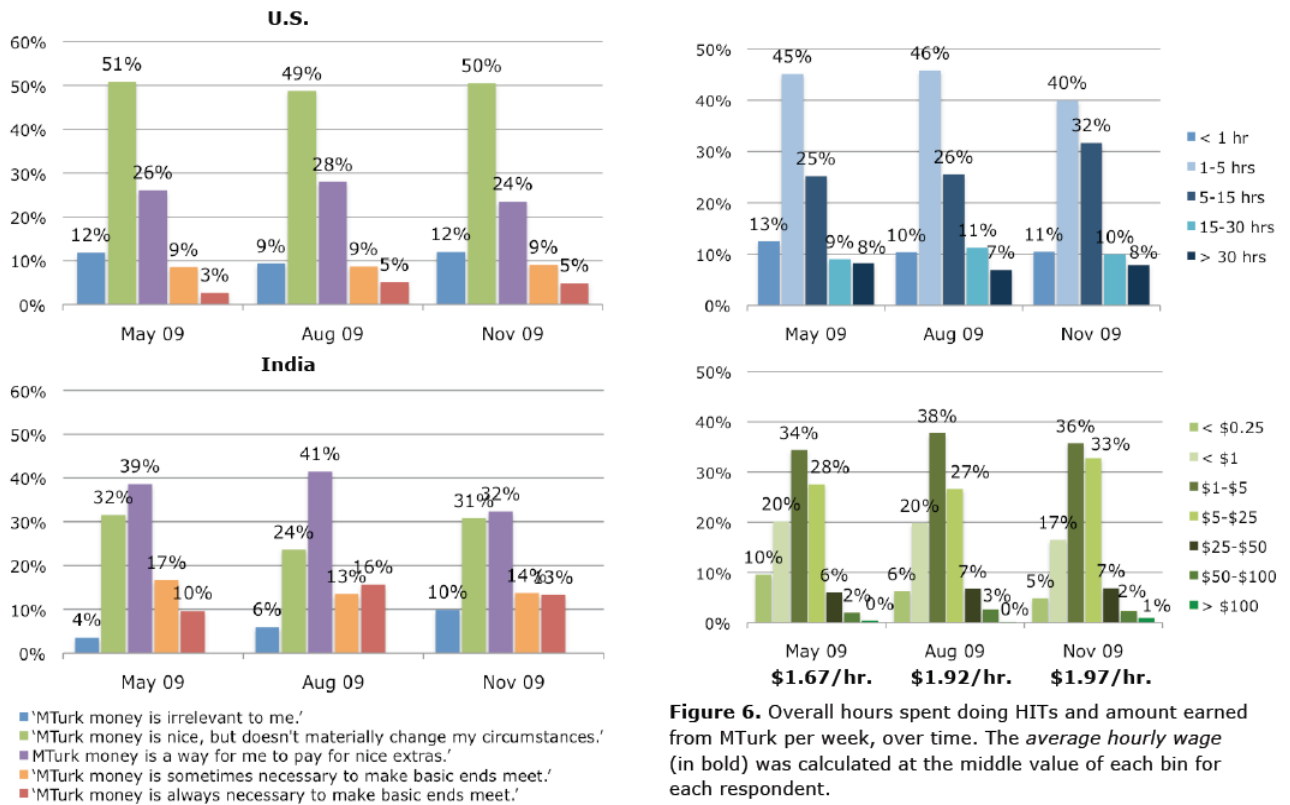
As it can be seen, Turker population is mainly composed by US and India users. While user from US have not increased significantly, people from India have grown considerably. Differences between US and India reside also in the gender of the Turkers: mostly female from US, mostly male from India.



The majority of users in the system are youngster with less than 34 years.



In spite of the low wages paid, the absence of motivation to deliver a better job, the actual impossibility to earn significant amount of money or improve user situation and income capacity, there is a significant amount of Indians who consider crowd sourcing as a full time job.



Amazon Mechanical Turk has been the most used crowd sourcing system in many studies and papers, such in example [15, 16, 19, 20].

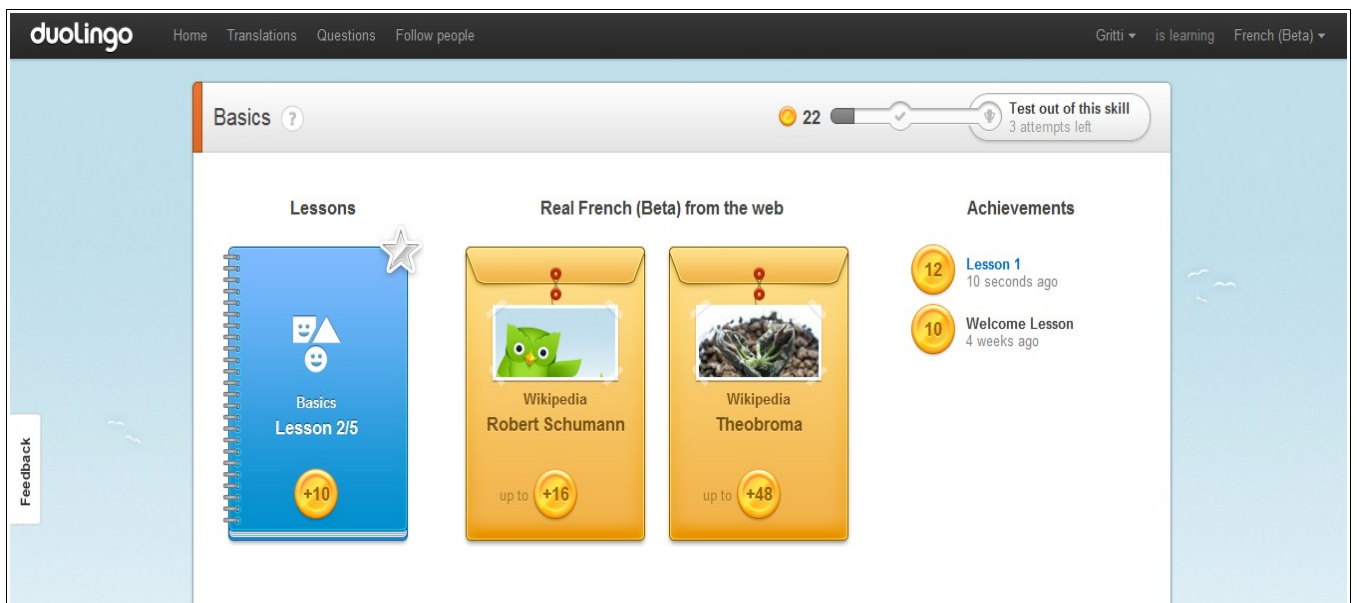
The name Amazon Mechanical Turk comes within an interesting anecdote: "The Turk". The Turk was a fake chess playing machine build in late 18<sup>th</sup>, and one of the most enduring hoax in history. The Turk was in fact a mechanical illusion that allowed a chess master to hide inside the machine and play instead of the machine. The secret of The Turk had last for several years, winning against other human opponents and creating a buzz into the audience.

This seems to be a proper name: The Turk was an example of an impressive system moved by a human engine.

### 3.4.2 Duolingo

*"Learn a language, while translating the web"*

Luis Von Ahn, co-inventor of the Captcha system <sup>\*</sup>, has been recently promoting Duolingo, a system thought to teach users new language and, at the same time, using their progress to translate sentences among the web.



The learning process offered by Duolingo uses a "listen and repeat" approach, where users are trained by looking at sentences already translated into target language by other users, and at any new class it keeps on adding new words, grammar rules and so on. The complexity of the sentences presented to user increases accordingly to user improvements. In example, a native English speaker that want to learn Spanish, he will start by looking at simple sentences such «donde esta la biblioteca?» and user can see the correspondent translation in English «where is the library?».

<sup>\*</sup>[\[en.wikipedia.org/wiki/Captcha\]](http://en.wikipedia.org/wiki/Captcha)

The ranking and improvements of users are represented with “experience coins”, that can be earned by doing courses or by providing translations, which are consequently verified by other users voting on the translation.

When user reach a certain amount of experience coins, he can begin to vote on other users translations, so to gain more coins.

Even if the system to learn new language that Duolingo propose is, in my opinion, limited under many aspects, this project has addressed some important points of a crowd sourcing project:

- Duolingo has clearly identified user's need: learn new languages. And it offers the reward clearly and, somehow, efficiently, in exchange of a limited service, which at the same time, help user in achieving his objective of becoming fluent in that language.
- Duolingo has clear direction, intention and it envisions the objective: it wants to translate every written text available in Internet, and it is starting with Wikipedia articles.
- Duolingo expressly target a defined set of people: all Duolingo users have the common objective to learn a new language.

Learning a foreign language is not an easy task, and Duolingo can be a useful tool to accomplish it. I found myself challenged from Duolingo many times, and I appreciated the learning method of increasing difficulties offered by the system, but I still consider the translation offered by a user that has just learned a new language, can not be that good.

In addition, one of the main pitfall of this project, is lack of cash rewards, which, at any rate, seems to be a key aspects in crowd sourcing. If just Duolingo would offer to users the possibility to learn new language, and consequently propose a system to make some money with the language just learned, participation in Duolingo would increase significantly.



### 3.4.3 Icanlocalize.com

I can localize offers professional service through its website\* where a community of professional translators can register and meet clients. Professionals translators have to apply to become members of Icanlocalize community, but they have to go through a certain process of verification before being accepted. During this verification process academia background has to be proven by uploading scans of the academia transcripts, and strict exams meant to verify the skills of the translator have to be passed. In other words, if a person is not a professional translator, nor own an academic degree in translation or foreign languages, he can not even register in the system.

This barrier limit the amount of users to professional translator, making Icanlocalize to be an online translation company, rather than a crowd sourcing system. Icanlocalize staff is in charge of all the management part, to avail applications to become users and, partially, to deal with clients. Translators have to behave at the same way as if they were hired from the company; i.e users have to inform the system about "planned leaves" such holiday and similar. Prices are really competitive though, and Icanlocalize has been used as a starting point to calibrate the reward offered in the designed system.

\*[\[www.icanlocalize.com\]](http://www.icanlocalize.com)

### 3.4.4 Crowdflower

CrowdFlower is a crowd sourcing service founded by Lukas Biewald and Chris Van Pelt, which made its public debut in 2009, and, since then, it has completed over 450 million tasks, and presently it is able to do 5 man-years of work in a day. Those numbers already give the idea of Crowdflower's importance in the crowd sourcing market. Differently to Amazon mechanical Turk, Crowdflower is meant to provide a range of enterprise solutions, especially for processing and creating large volume of data. Crowdflower also provide, at some extension, features for project management, helping the requester in dividing a project in smaller task and distribute them to the crowd workforce. Crowdflower also provides a quality management system called Gold Standard Data, where some contributors perform some pre-completed task with the objective to determine the accuracy and the trustworthiness of the worker (rather than the work provided itself).

**CrowdFlower** Products Contributors Company Blog Docs

# Products

🔍 How Crowdflower Works

- RTFM**  
Real Time Foto Moderator
- Senti**  
Sentiment Analysis
- Ranked**  
Search relevance tuning
- Categorize**  
Products, Businesses, & More
- Content Generation**  
Instantly create quality SEO content at scale
- Data Collection**  
Extract meaningful data from websites
- Data Verification**  
Increase the accuracy of your data
- Surveys**  
Get results quickly

CrowdFlower is not exactly a crowd sourcing system, but it is a crowd sourcing aggregator. In other words, it provides an management interface to use properly others crowd sourcing platform such as Amazon Mechanical Turk.

The screenshot shows the Amazon Mechanical Turk interface. At the top, there's a navigation bar with 'Your Account', 'HITS', and 'Qualifications' tabs. A notification indicates '301,636 HITS available now'. The user 'andrea gritti' is logged in, with links for 'Account Settings', 'Sign Out', and 'Help'. Below the navigation, there are filters for 'All HITS', 'HITS Available To You', and 'HITS Assigned To You'. A search bar shows 'Find HITS containing crowdflower that pay at least \$ 0.00'. There are checkboxes for 'for which you are qualified' and 'require Master Qualification'. The main content area is titled 'HITS containing 'crowdflower'' and shows '1-10 of 67 Results'. The results are sorted by 'HITS Available (most first)'. There are five HIT listings, each with a title, requester, expiration date, time allotted, reward, and number of hits available. The first listing is 'To Which Category does this Product Belong? (1/20)' with a reward of \$0.05 and 9976 hits available. The second is 'FIND AN ADDRESS ON A WEBSITE! (International)' with a reward of \$0.04 and 9876 hits available. The third is 'Search Google and answer questions about the search results' with a reward of \$0.05 and 2121 hits available. The fourth is 'Categorize addresses as Residential (Homes) or Commercial (Businesses)' with a reward of \$0.04 and 1977 hits available. The fifth is 'Find Fun Activities in or near a City' with a reward of \$0.10 and 1901 hits available.

HIT Title	Requester	HIT Expiration Date	Time Allotted	Reward	HITS Available
To Which Category does this Product Belong? (1/20)	CrowdFlower	Jul 26, 2012 (6 days 23 hours)	60 minutes	\$0.05	9976
FIND AN ADDRESS ON A WEBSITE! (International)	CrowdFlower	Jul 26, 2012 (6 days 22 hours)	60 minutes	\$0.04	9876
Search Google and answer questions about the search results	CrowdFlower	Jul 26, 2012 (6 days 23 hours)	60 minutes	\$0.05	2121
Categorize addresses as Residential (Homes) or Commercial (Businesses)	CrowdFlower	Jul 26, 2012 (6 days 23 hours)	60 minutes	\$0.04	1977
Find Fun Activities in or near a City	CrowdFlower	Jul 26, 2012 (6 days 19 hours)	60 minutes	\$0.10	1901

### 3.4.5 InboxDollar

InboxDollar is a mistrustful system. Meaning that it can be used legitimately, but it has a bad reputation in Internet, and here there is why.

At registration time, InboxDollar profiles user in quite detailed manner, and so user already gets the sensation that inboxdollar may be a scam.

## Interests Survey

Allow us to customize your experience at InboxDollars by completing the following survey:

For which of the following areas do you have an interest? (Please select as many as apply)

- |   |  |
|---|--|
| <input type="checkbox"/> Books and Magazines          | <input type="checkbox"/> Health and Fitness      |
| <input type="checkbox"/> Business and Entrepreneurism | <input type="checkbox"/> Music and Entertainment |
| <input type="checkbox"/> Computers and Technology     | <input type="checkbox"/> Shopping                |
| <input type="checkbox"/> Arts, Crafts and Hobbies     | <input type="checkbox"/> Travel                  |
| <input type="checkbox"/> Cars and Trucks              | <input type="checkbox"/> Sports                  |
| <input type="checkbox"/> Family and Home              | <input type="checkbox"/> Pets                    |
| <input type="checkbox"/> Finance and Investments      | <input type="checkbox"/> None of the above       |

What types of magazines do you enjoy? (Please select as many as apply)

- |  |  |
|--|--|
| <input type="checkbox"/> Automotive              | <input type="checkbox"/> Food/Cooking                |
| <input type="checkbox"/> Computers               | <input type="checkbox"/> Health                      |
| <input type="checkbox"/> Fashion                 | <input type="checkbox"/> Men's Issues                |
| <input type="checkbox"/> Home/Home Services      | <input type="checkbox"/> Music/Entertainment         |
| <input type="checkbox"/> News/General Interests  | <input type="checkbox"/> Outdoors                    |
| <input type="checkbox"/> Sports/Recreation       | <input type="checkbox"/> Science/Technology          |
| <input type="checkbox"/> Business/Industry       | <input type="checkbox"/> Travel/Leisure              |
| <input type="checkbox"/> Children/Baby/Parenting | <input type="checkbox"/> Women's Issues              |
| <input type="checkbox"/> Crafts/Hobbies          | <input type="checkbox"/> Other                       |
| <input type="checkbox"/> Ethnic/Cultural         | <input type="checkbox"/> Not Interested in Magazines |

In which of the following activities do you participate? (Please check as many as apply)

- |  |  |
|--|--|
| <input type="checkbox"/> Attending Concerts        | <input type="checkbox"/> Home Decorating   |
| <input type="checkbox"/> Cooking                   | <input type="checkbox"/> Entertaining      |
| <input type="checkbox"/> Attending Sporting Events | <input type="checkbox"/> Gardening         |
| <input type="checkbox"/> Dining Out                | <input type="checkbox"/> Camping           |
| <input type="checkbox"/> Bicycling                 | <input type="checkbox"/> Fishing           |
| <input type="checkbox"/> Jogging                   | <input type="checkbox"/> Hiking            |
| <input type="checkbox"/> Tennis                    | <input type="checkbox"/> Reading           |
| <input type="checkbox"/> Walking                   | <input type="checkbox"/> Photography       |
| <input type="checkbox"/> Collecting/Scrapbooking   | <input type="checkbox"/> None of the above |

Next Page

Previous Page

With few more research in internet, it resulted that Inboxdollars offers the possibility to earn some money, but it does not mean users necessarily will earn money.

Users are asked to do tasks such receiving email advertisements for various programs or products and then answer the relative survey. Aside from receiving spam, users can also earn money for other tasks such play games, sign up for other users surveys, sign up for programs, shop, and search on internet.

Looking into few of the programs that are heavily promoted throughout Inboxdollars, there is, for example, eBay. InboxDollars say it pays \$6.00 to sign up at eBay and place a bid, and Inboxdollars makes anywhere between \$25-\$35 every time a new user signs up at eBay and places a bid. Another example is Stamps.com which is also heavily promoted on Inboxdollar, where users can earn \$10 bonus just for signing up for their free trial. However, if the user forget to cancel the free trial after a couple of weeks, he will be billed \$15.99 per month and InboxDollars makes \$50.

This system is just an example how crowd sourcing systems can also be used to fool users.

## 4 Design

This chapter describes the main objective of the project: the design of a crowd sourcing system for Translation and Software localization. A couple of versions of the designed system are described, so to present the reader the process that lead to the final version of the architecture.

The Design of the architecture started in February 2012 and lasted until April 2012. During those months different versions of the architecture have been proposed, until a final version has been judged satisfactory by the team, and it was presented to CA Technologies legal department to be analyzed as it may be patentable.

From the study of the current state of art described in the previous chapter, it was clear the lack of quality management system is limiting crowd sourcing potential diffusion and adoption. The collaboration of a massive amount of people who are geographically distributed it seems to be a barrier for quality. In general, it is rather complicated to define automatic mechanisms in order to monitor quality, because the lack of automatic methods to solve the kind of problems which suit best crowd sourcing usually implies a raw definition of the concept of quality, and consequently a high complexity in order to establish a strong quality measure and control.

Therefore, the main idea that drove the entire design phase it had been the objective of adding a quality management layer above the crowd.

As per requirements, the quality provided by the system should not entirely rely on the crowd because, as it has been described in the previous chapters, quality in crowd sourcing system is aleatory.

Once again: crowd stands for an anonymous group of people able to provide a certain result on different tasks; the quality of the result totally depends on the will and skill of the worker. Thence, the designed system output's quality must

be guaranteed by the system itself, indifferently of the skillsets of the workers involved, nor the total amount of users needed to complete a specific process or task.

In detail, the design phase focused on the creation of an innovative and robust architecture which present the following novel features:

- A quality management layer above the crowd able to provide an excellent result in translation;
- Ranking system for users, in order to keep track and define users skills and behavior in the system;
- Efficient resources (workers) management, in order to valorize better translators among a crowd of untrusted and unknown users;
- Adaptive work process able to readjust and harmonize the amount of steps needed to provide quality in translation. The work process should be modified according to the skill and rank of the worker selected for the process.

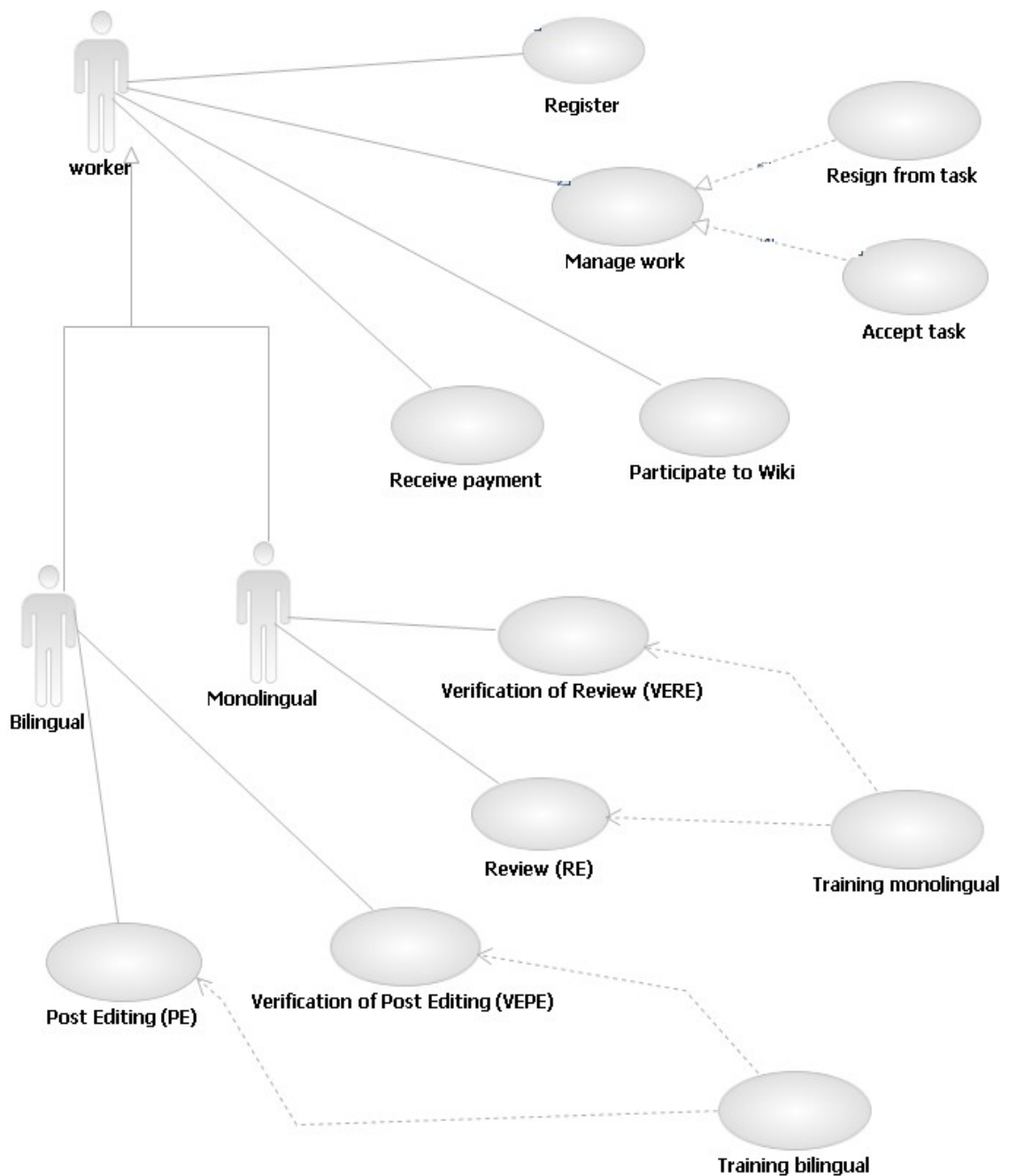
To recapitulate, the designed system is meant to orchestrate resources inside a crowd made by unknown and untrusted users, each one with different skills and capability in translation, and consequently reduce and adapt the translations phases needed to guarantee a certain level of quality in the translation, accordingly to the ranking of the users involved in the process itself.

The designed architecture has also objective to reduce the amount of workers involved in the process, hence to reduce the amount of money needed for a translation, while still assuring quality in the result.

## **4.1 Use case**

Use cases represent the utilization of the system, and they are useful to formalize the different way that users interact with the system. In the following image the main use cases of the system are listed.





In the system there are mainly two kind of users: monolingual users are people that can only speak a single language, while bilingual users are people that know two or more languages. Users have common use cases and different use cases, accordingly to the language user knows. Use cases are better described in the following pages.

Use Case	Register	
What	User register in the system	
Who	worker	
Main Case		
Actor	System	
1-User connect to the web page, and surf to the register page	2-System show the login page and the Captha	
3-User fullfill the registration form, resolve the captcha and press the submit button	4-System save user data and send email to user's address confirming the registration	

The user is can register by filling in his/her basic personal data to register and log in to the system. A captcha (Mollom service) should be used to verify that it such a registration is not an automatic registration. These captchas will not be utterly complex for the user (ex: ReCaptcha). User also should receive a confirmation email address and a short guide on how to start to use the system.

## Manage work

Once user is registered, he can manage the different tasks the system can propose him; basically he can accept a new task or resigns from a task user has assigned. User can resign from assigned task if he feel he can not complete it, and resign from a task influences less negatively his ranking, than missing a deadline for a task.

Use Case What Who	Resign from a task User resign from a task he had previously got assigned worker
Main Case	
Actor	System
1-User decide he can not finish the task before the deadline. He logs in the system  3-User select the task and choose the option resign	2-System show the user's personal page where the task he has assigned are listed.  4-System take off the task from user list; system update user ranking; System propose the task to another user

System proposes to user just the tasks he can handle (i.e: if user inserted at registration time he knows Spanish native speaker, Italian tasks not offered to ). When system propose the task, users has a deadline to accept the task, and a deadline to deliver the task.

Use Case	Accept a task	
What	User accept a task proposed by the system	
Who	worker	
Main Case		
Actor	System	
2-User receive email, read the deadlines and the description of the task, accept or refuse the task by clicking the link in the email	1-System propose a task to user via email  3-If system receive a positive answer, system assigns task to user. If the deadline to accept the task expire, the task is proposed to another user.	

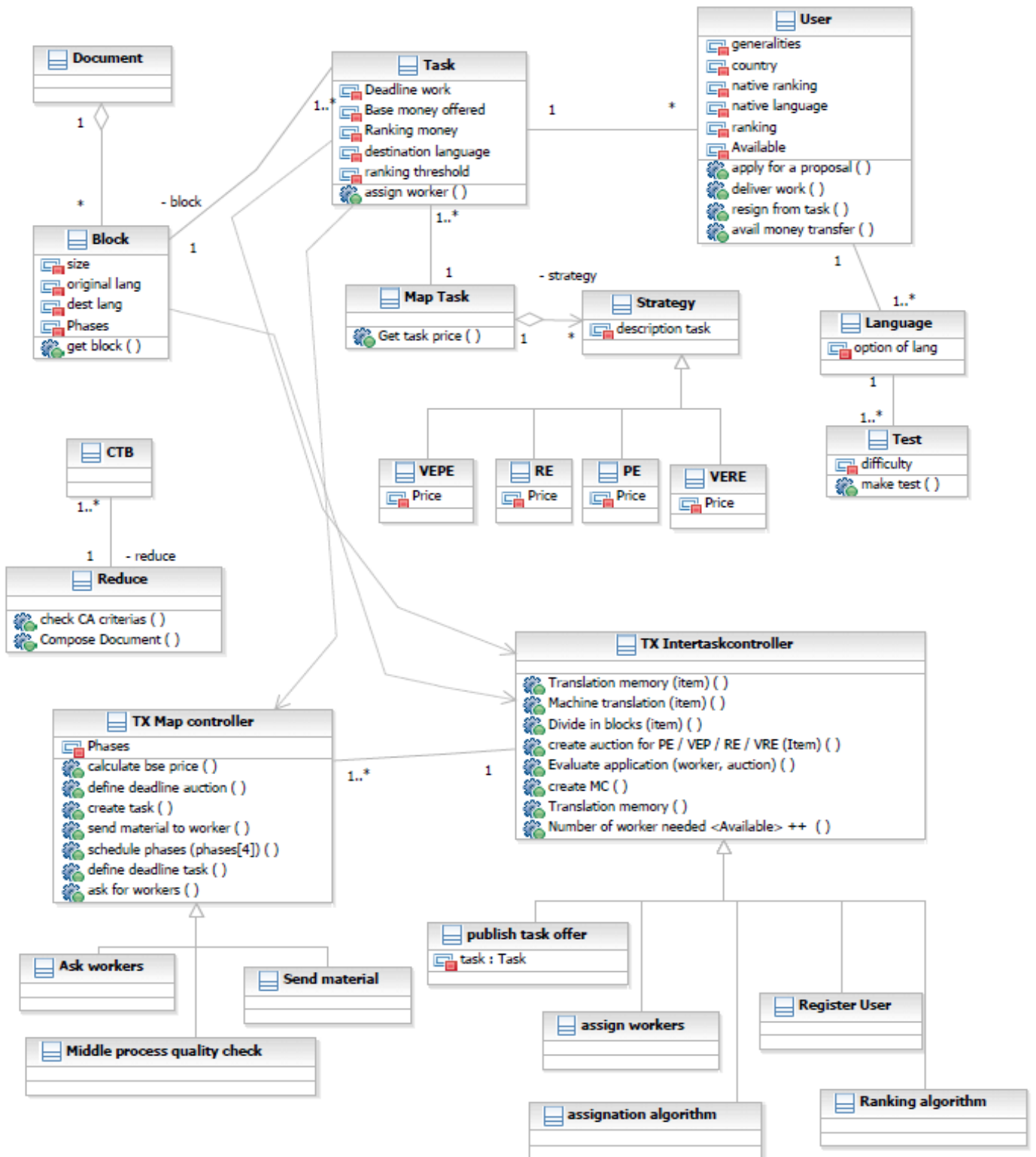
Use Case	Receive money	
What	User avail the money transfer to his bank accoun	
Who	worker	
Main Case		
Actor	System	
1-User log in	2- System show user personal page	
3-User avail money transfer	4-System check if the amount of money earner from the user is enough to avail the money transfer	
	5-In case previous answer is positive, money transfer is started, and the amount of money earned by user is reset.	

Receive Payment: It has been decided that User has to avail the money transfer, so to avoid the few money each user earn for each task.

Participate in the forum: user can participate to wikis, forums and different blogs. This has been thought in order to create a community, and exploit the fact that crowd wisdom is more effective when user feels to be part of a community of users.

The rest of the main use cases represent the behavior of the system, so they are described in detail the following chapters.

## 4.2 Logical View



(Logical view of the designed architecture)

## 4.3 First Version

In this chapter it is described the first version of complete architecture which was presented to CA Labs. The first complete version was proposed in the end of March 2012 after one month of the beginning of the project. The description provided in this chapter shows the overall behavior of the designed system, while a more detailed description is provided in chapter 4.4.

### 4.3.1 Description of the system

The first version of the architecture is composed by 3 main modules:

- 1) Task controller **(TC)**
- 2) Inter-task controller **(ITC)**
- 3) Worker ranking module **(WRM)**

Those 3 modules are in charge of orchestrating the translation process of the text.

The system receives as input any kind of text in the original language, and it is set the target language to be translated to. When the text enters the system, TC takes it and performs a first translation with the latest technology in Machine Translation. When the Machine translation is completed, TC branches the translated text in smaller chunks of a defined size. Each chunk is then organized according to MapReduce paradigm [24]. The Map Reduce programming model has been chosen because translating a text represents a problem which can be easily separated in smaller problems. Since there is no dependency among the different chunks in which the initial text is divided into, the translation of each chunk can progress separately, de facto representing an

embarrassingly parallel workload. Hence, MapReduce seemed to be the most suitable programming model for processing the translation of a text.

When the initial text has been divided and organized, TC requests workers to ITC.

ITC is the module responsible of managing users and allocating workers to each chunk of text. At any time in the system, the amount of TC loaded in memory corresponds to the number of text to be translated, while ITC is unique in the entire system and interacts with all TC at the same time, managing the overall distribution of workers. ITC keeps the list of free workers and assigns workers to different TC, according to the languages workers are able to work on.

When TC receives the workers, the translation process starts for each chunk of the text. **The translation process is divided in 4 basic steps:**



**Step 1–Post edit:** The chunk of the initial text is assigned to a worker for post edit the machine translated output. The worker is given the text in the original language and the machine translated text to post edit.

Use Case	Post Edit	
What	User improve the Machine translate text	
Who	Bilingua worker	
Main Case		
Actor	System	
1-User log in	2- System show user personal page with all the task assigned to him	
3-User start to work on a Post Edit task	4-System show the Original text in the original language, and the machine translated text	
5-User modify the machine translate text and submit the task when he feel satisfied	6-System save the user work	

**Step 2–Verification:** The output from step 1 is given as input in step 2. Post edited text is then passed to a group of Verificators. Verificators are in charge of looking at the post edited text and find errors ( typos, grammar errors, and so on ). If the error rate found from all the Verificators is too high, the text is stopped, no further action is taken and a request for re-execution of the partition is made. The error rate is also reported to the WRM. If the error rate is not too high, the chunk of text continues to the 3<sup>rd</sup> step.

Use Case	Verification of Post Edit	
What	User verify the work done at step 1	
Who	Bilingua worker	
Main Case		
Actor	System	
1-User log in	2- System show user personal page with all the task assigned to him	
3-User start to work on a Verification of Post Edit task	4-System show the Original text in the original language, the machine translated text, and a set of error among which user can select	
5-User select error in the text and submit the task when he feel satisfied	6-System wait untill all the Verificators have completed the task in order to decide what to do. System increase the user money of the price paid for the task	

**Step 3–Review** : In this step a native speaker of the target language is presented with the text translated in the previous 2 steps. User assigned to this step has to improve the fluency and the structural coherency of the text. Since user is working with his native language, he supposedly can improve the text.

Use Case	Review	
What	User improve the work done at previous step	
Who	Monolingual worker	
Main Case		
Actor	System	
1-User log in	2- System show user personal page with all the task assigned to him	
3-User start to work on a Review task	4-System show the Post Edited text	
5-User improve the general quality of the text in his native language and submit the task when he feel satisfied	6-System save the text which have been reviewed.	

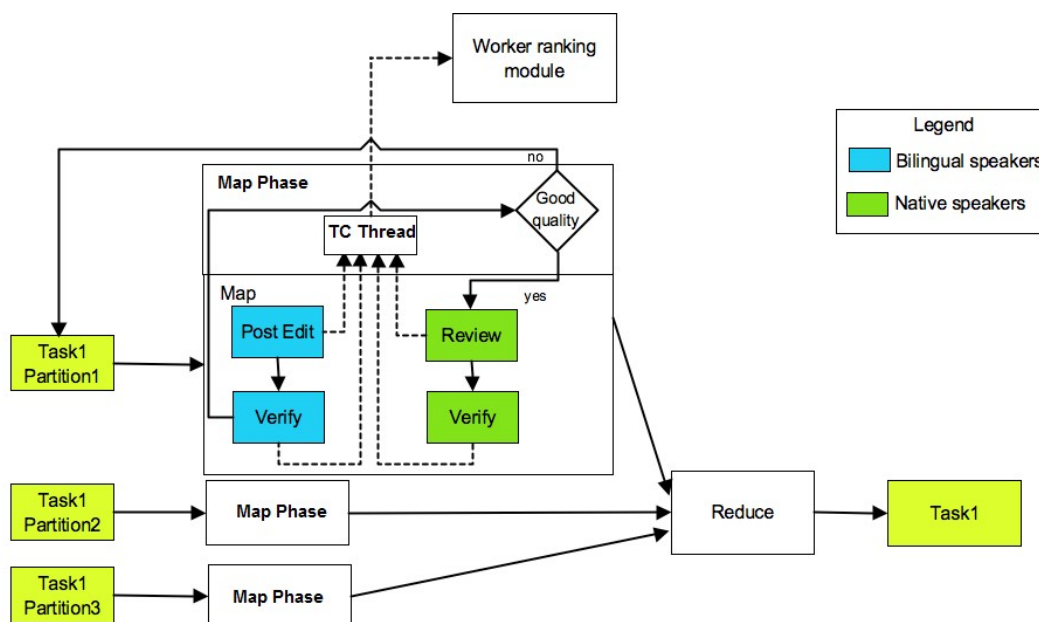
**Step 4– Verification of Review :** The reviewed text is received by another set of native speaker workers who are in charge to check if no errors were introduced during the Review step. Same as in step 2, the chunk of text continues in the translation process if the error rate is not too high.

Use Case	Verification of Review	
What	User verify the quality of the reviewed text	
Who	Monolingual worker	
Main Case		
Actor	System	
1-User log in	2- System show user personal page with all the task assigned to him	
3-User start to work on a Verification of Review task	4-System show the improved text which is the ouput of step 3, within a set of possible error that worker can select among	
5-User find all the errors in the text and submit the task when he feel satisfied	6-System wait untill all the Verificators have completed the task in order to decide what to do. System increase the user money of the price paid for the task	

At any step, if provided, glossary and style guide are given to users.

As it can be seen, bilingual speakers are assigned to work where the original language and the target language are involved (step 1 and step 2), while monolingual speakers are assigned to work on text where only their mother tongue is needed (step 3 and step4), making the system available for any kind of user. Those 4 basic steps are formalized in the paper [25] as an Action-Verification Unit. The Action Verification Unit is better explained in chapter 4.7

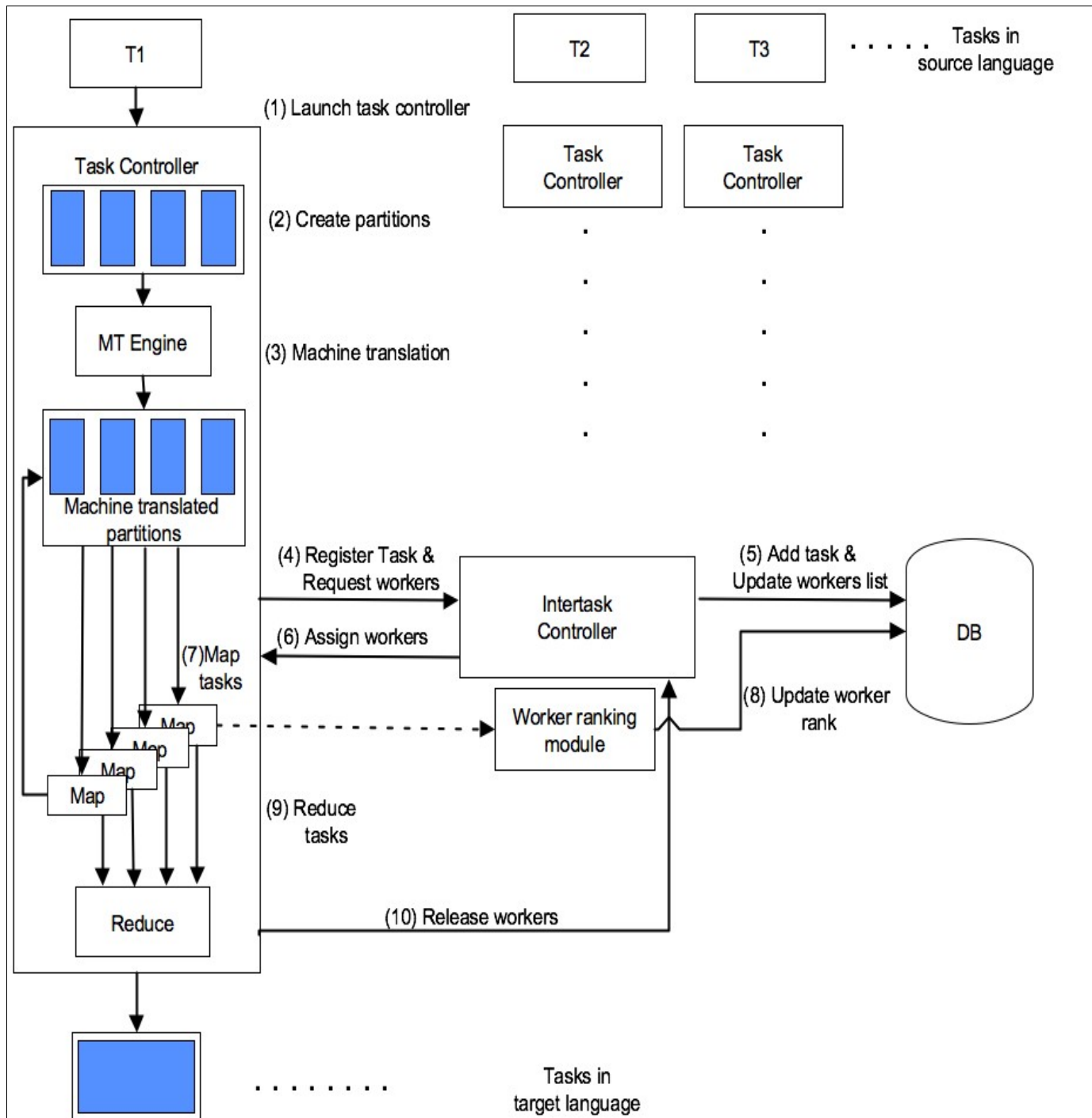
When all parts of a same text have been translated correctly, the Reduce phase recollect all texts that belong to a single document, and join together all the chunks in order to create the final translated text. As a reminder, there is one TC for each different text that is being translated.



When all the different steps of the above mentioned translation process are over, the TC also informs the WRM about the workers performance, reporting if workers at Post edit(PE) and Review(RE) phases have done a good or bad job accordingly to the error rate reported by the workers in the verification steps, and so WRM calculates a new rankings for those workers

As a final step, each TC releases the workers it received by the ITC, and ITC sets those workers as available, so they can be selected to work on other tasks.

## Overall view of the System architecture



(Designed Architecture V1)

#### **4.4 Final Version**

The design phase has been a continuous improvement, where each meeting with CA Labs was used to calibrate the direction and adjust the prerequisite analysis. The system has been evolving through different versions, and this chapter describes the final version of the architecture which has been proposed at the end of March 2012, after 2 months of studies, and it was fully accepted by CA Labs.

The design phase have been focusing on many different problems, such, in example:

- In phase 3, Reviewer could change the meaning of a Post edited text, and no one could realize it, since workers at step 4 were not provided with the original text that worker at Review phase improved;
- There were no control on the Verificator work. Thus, their ranking was not influenced by the work done at verification phase;
- All the knowledge produced by user was lost; Translation memory was not kept;
- When TC was splitting the text in different chunks, it was then asking for all the workers it needed, and, if the text was long, too many workers were allocated at the same time. This was making TC to static, and a possible bottleneck too;
- The wisdom of the crowd is more efficient if the user feels a sense of community, so system needed tools such blogs, forums, wikis and so on;
- According to experience of the Translation and localization department of CA Technologies, the translation process is not something static, and it is

adaptable to the skills of the translator. In other words, not all the steps are always done.

The final version of the designed system is more complex, and it has more features compared to the version described in the previous chapter. The final version has also moved the objective from managing the quality of the work provided by an user, to a profiling of the behavior of the user. Instead of just looking at the quality of a specific work provided in a specific phase, profiling user has become more important because, in the final version, the designed architecture relies on the idea that if a user has been providing good quality work until now, it is likely that he will provide good quality work also the next time. Consequently, the final version of the designed system is able to organize the translation process putting in place less quality checks for a trusted user, rather than keeping the entire translation process static for all the workers.

Therefore, the ITC makes larger use of the ranking system, making the ranking to be the cornerstone where to calibrate all the translation process, and consequently distribute users to different tasks in the best possible manner.

From here, the final version of the architecture includes also a new module that allow a finer granularity, and shifts the main focus of the system from the translation of an entire text to the translation of chunk, resulting in a more flexible and elastic system. The final version of the system allows a user to work on different chunks of the same text, improving consequently the coherency and the style of the final product.



#### 4.4.1 Description of the system

The final version of the architecture included the following modules:

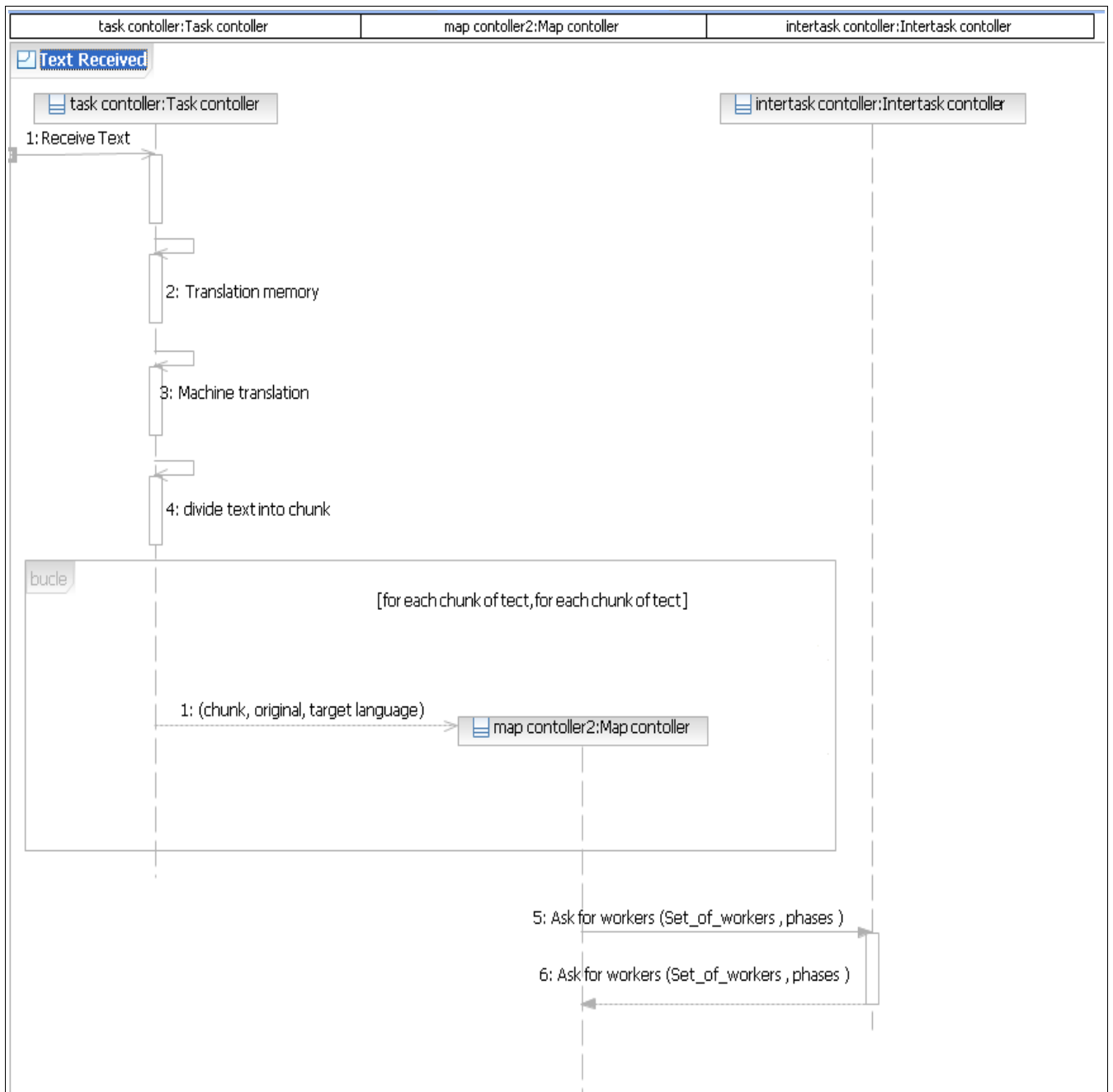
- 1) Task controller (TC)
- 2) Inter-task controller (ITC)
- 3) Worker ranking module (WRM)
- 4) Map Controller (MC)

The system flow is quite similar to the previous version, and the main difference is the design of Map Controller, which has been created to allow a finer granularity. As a consequence, Task Controller module is not the responsible of the translation of each chunk anymore, but instead it works as a coordinator for all the MC, to which is assigned a map phase. TC is still responsible of the initial machine translation and the division of the input text.

Moreover, the final version is able to improve the automatic translation by keeping a translation history of the previous translations, so to do not lose previous work and knowledge created from the crowd, since it often happens that new documents to be translated are just older version of the same document, which has been updated.

## Task Controller

The input text is received by Task Controller (TC). The text is matched with translation memory (corpus of previously translated texts) and if exact match of a statement is found, it is translated into the target language. The remaining text goes to a machine translation engine that performs automatic translation. Again, the machine-translated texts are divided into different blocks and distributed employing MapReduce process using the crowd. Once the text has been divided in chunks and the MapReduce has started, TC creates a MC for each chunk of text.

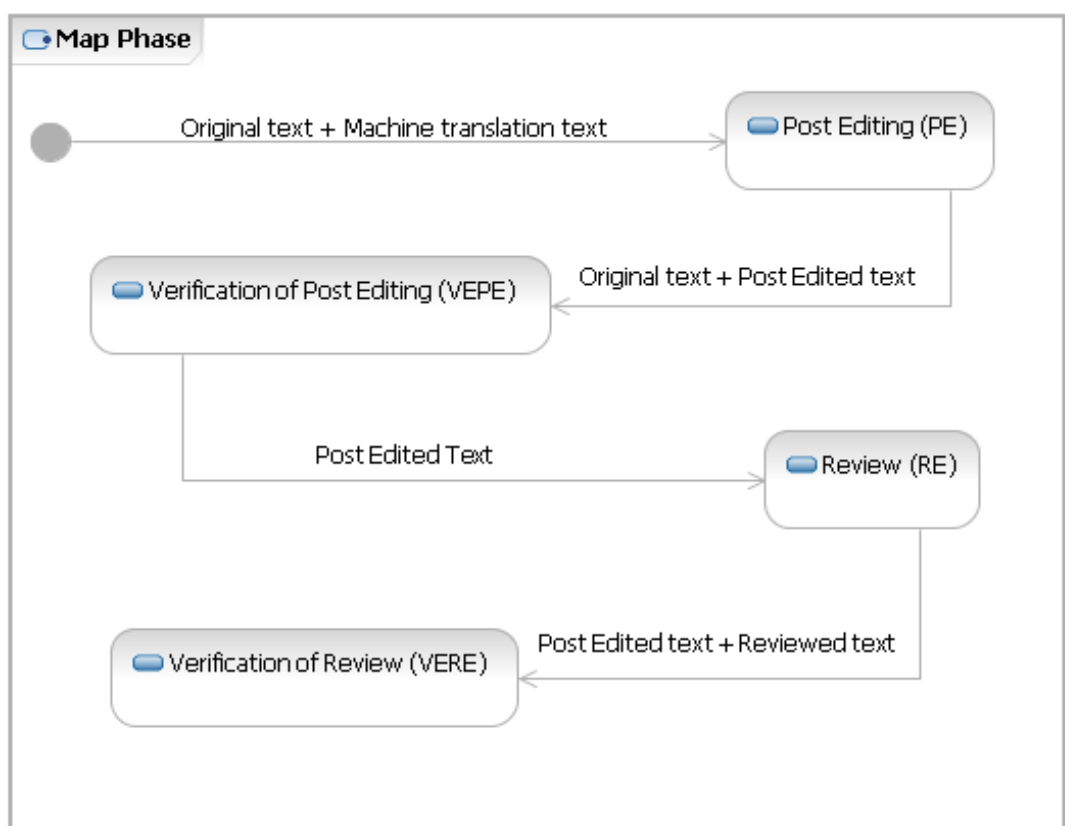


## Map Controller (MC)

Map Controller is the module responsible for the successful completion of the translation process of a chunk of text. It handles the different phases and it is the one who interacts with the Inter-Task Controller (ITC) for workers.

Map Controller is the module that orchestrates the different phases that have to be performed in the translation process. MC does not decide which of the phases will be performed, but it receives this as input from the ITC.

Similarly to the version described in the previous chapter, the main steps are 4.



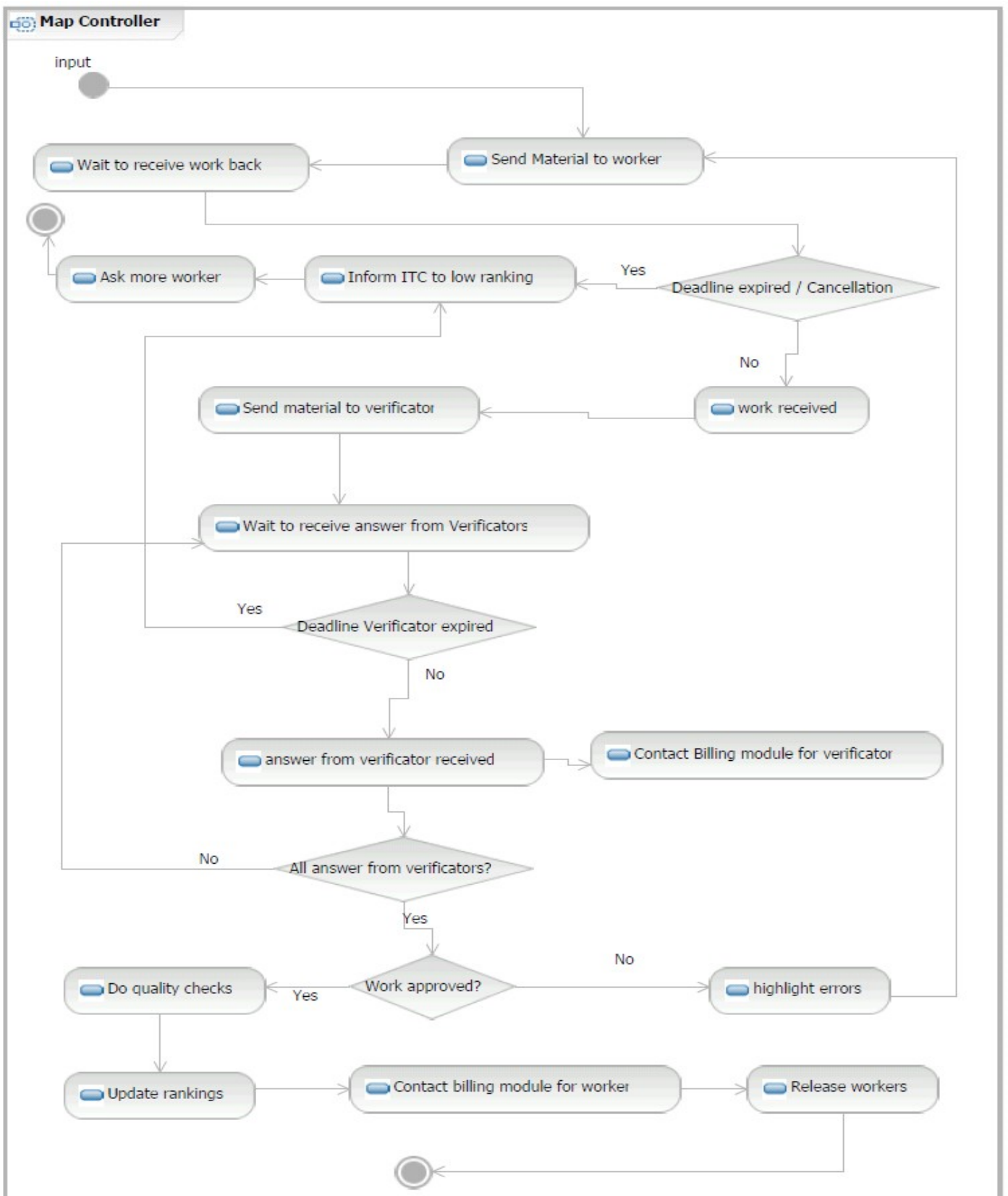
Likewise, steps 1 and 2 are performed by bilingual speakers of source and target language, and Steps 3 and 4 are performed by native speakers of target language. In Step 1, the partition is assigned to a worker from the crowd, for post edit. The worker is given the original text in source language and the machine-translated text. In Step 2, the partition which has been post edited is sent to one or more workers for verification, and the text in original language is

sent along with it, so the workers at verification phase can check if the meaning of the original text has been correctly translated by the first worker.

These workers find mistakes in translation, and report them to the Worker Ranking Module. If the error rate is too high, no further steps are carried out and a request is sent for re-execution of Map process for this partition. If the error rate is low, the text with highlighted errors is sent back to the post editor of Step 1 for revision, and this phase is called Post edit phase 2. At phase Post edit 2, the same worker can correct the errors found by Verifiers, and provide a final version which continues to the next phase.

After Post Edit phase, the text is sent to Step 3. In Step 3, the native speaker of the target language improves the quality of the translation, ameliorating grammar, structural problems and other blunders, and corrects them. In Step 4, the reviewed text from Step 3 along with the text from Step 2, is sent to one or more native speakers for verification, to check if there are still any errors in the language, the structure or flow of the text. The text from the Post Edit phase is also sent to Verification of Review phase, in order to allow workers to check that the meaning was not changed during the Review phase.

The error rate is again reported to the Worker Ranking Module. The worker-ranking algorithm assigns ranks to the workers based on the error rate reported by the verification step. The Worker Ranking Module updates the rank of each worker after the completion of every job. If a worker does not complete the job in the given time, or the overall quality of work is too low, the partition is re-executed. Finally the partition is sent to the Reduce process, which is responsible for combining all the partitions and producing the final item in target language. After each map phase, the MC releases all the workers by notifying the ITC who updates the list of available workers. The final product is fed into the translation memory for future use. After the completion of each phase, the MC sends a notification to the billing module, which increases user's credits by the amount paid for the task user performed.



### **Inter-Task Controller (ITC)**

ITC is in charge of managing users. When ITC receives a request for workers from different MC, it selects workers according to the language needed by the MC.

For step 1 Post edit user is selected among the group of users that have set their status as available. When users set their status as available to receive tasks, it means that the system can select them and assign them a task.

For step 2: Verification of Post Edit, and step 4: Verification of Review the workers are assigned to the task in a balanced manner by the ITC, in order to create groups of workers where the total amount of their ranking is equilibrated all over the groups of workers assigned to the different MC: the higher the rank of the available workers, the smaller the number of workers are assigned, and vice versa.

In addition, ITC also decides how many phases the MC should perform, according to the rank of the selected workers.

In other words, Map Control has to manage the translations steps which can vary according to the ranking of the users involved. When system proposes a task to the worker, it informs him about the acceptance deadline and the delivery deadline. If worker accepts, he is assigned to the task. If the user selected for the Post edit phase has really high rank, ITC may decide that the other phases are not needed. Conversely, if the worker selected at Post edit phase has low rank, ITC decides how many workers are needed at Verification phase, in order to obtain high quality for the final translation. Once again, the amount of workers chosen for Verification phase strongly depends on their ranking: if many users with low rank accept the task, there will be assigned more of them, if instead users with high ranking are available, just few of them are assigned to the Verification phase.

Previous architecture was assigning users in a more static manner, and also the new phases of Post edit 2 and Review 2 did not exist, since it was supposed that errors found by Verificators may be modified and correct in the following

phases. Since in the final version of the designed architecture, the translation process changes according to the skills of the workers involved, it may be that just one single phase is done.

Summarizing, not all the phases are included in the translation process whenever the ranking of the worker at the Post edit phase is high enough, and the system strongly believe that the work done at this phase is already of good quality.

The following table represents the probability of each phase to happen. The probability are totally made, and they should be tuned by running some experiments where to check how quality in translation depends on the different combinations of phases and user's ranking.

<b>Post Editor Rank</b>	<b>Prob of Verification</b>	<b>Prob of Reviewing</b>	<b>Reviewer Rank</b>	<b>Prob of Verification</b>	<b>Cost</b>
>90	0.2	0.4	>90	0.2	$2+(1*0.2)+(0.75*0.4)+(0.75*0.2*0.2) = \mathbf{2.53}$
80 to 90	0.7	0.6	80 to 90	0.6	$1.5+(1*0.7)+(0.75*0.6)+(0.75*0.6*0.6) = \mathbf{2.92}$
60 to 80	1	1	60 to 80	0.8	$1+(1*1)+(0.75*1)+(0.75*0.8) = \mathbf{3.35}$
40 to 60	1	1	40 to 60	1	$0.5+(1*1)+(0.75*1)+(0.75*1) = \mathbf{3.00}$
<40	-	-	<40	-	-

To motivate users to deliver a better job, thus increasing the quality, the money paid for each task is the sum of a base price which only depends on the size of the block and the task that has to be performed on it, and an extra amount of money which only depends on the ranking of the worker.

ITC is in charge of assigning users to the different MC, and also to decide the phases accordingly to the ranking of the user. ITC can assigns users in the best manner in order to reach a certain level of quality, based on the ranking of the users, but, if the main objective is to control the cost of each translation, ITC can select users according to a cost algorithm in a manner that a translation

cost stay within a fixed amount of money.

The algorithm works like this: suppose it is allocated a budget of 'x' cent per word for the verification task. The system will start selecting workers from the queue of available users and add their cost to the total cost. When the total cost equals the allocated budget, the selection is stopped. if at any time during this process the cost exceeds the budget, that worker is skipped until a next worker with lower cost is found that satisfies the budget. The pseudo code would be something like the following:

```
pick worker from queue;
while(total_cost != allocated_budget)
{
    if (total_cost + worker_cost > allocated_budget)
    {
        skip to next worker;
    }
    else
    {
        select worker;
        total_cost = total_cost + worker_cost;
    }
}
```

The total\_cost is the base price paid for the task (that depends on the amount of words and the task that has to be performed), while worker\_cost is the extra money which is paid accordingly to the ranking of the user.



In any case, ITC assignment algorithm assures the following:

- Post Editor worker can not be assigned to Verification of PE for the same chunk of text, neither he can be assigned to Reviewing.
- Reviewer worker can not be assigned to Verification of RE for the same block, neither he can be the same user assigned to Post edit phase;
- Worker can be assigned to only one block. when a worker is assigned, he is labeled as busy and he can not get assigned to others chunk of text.

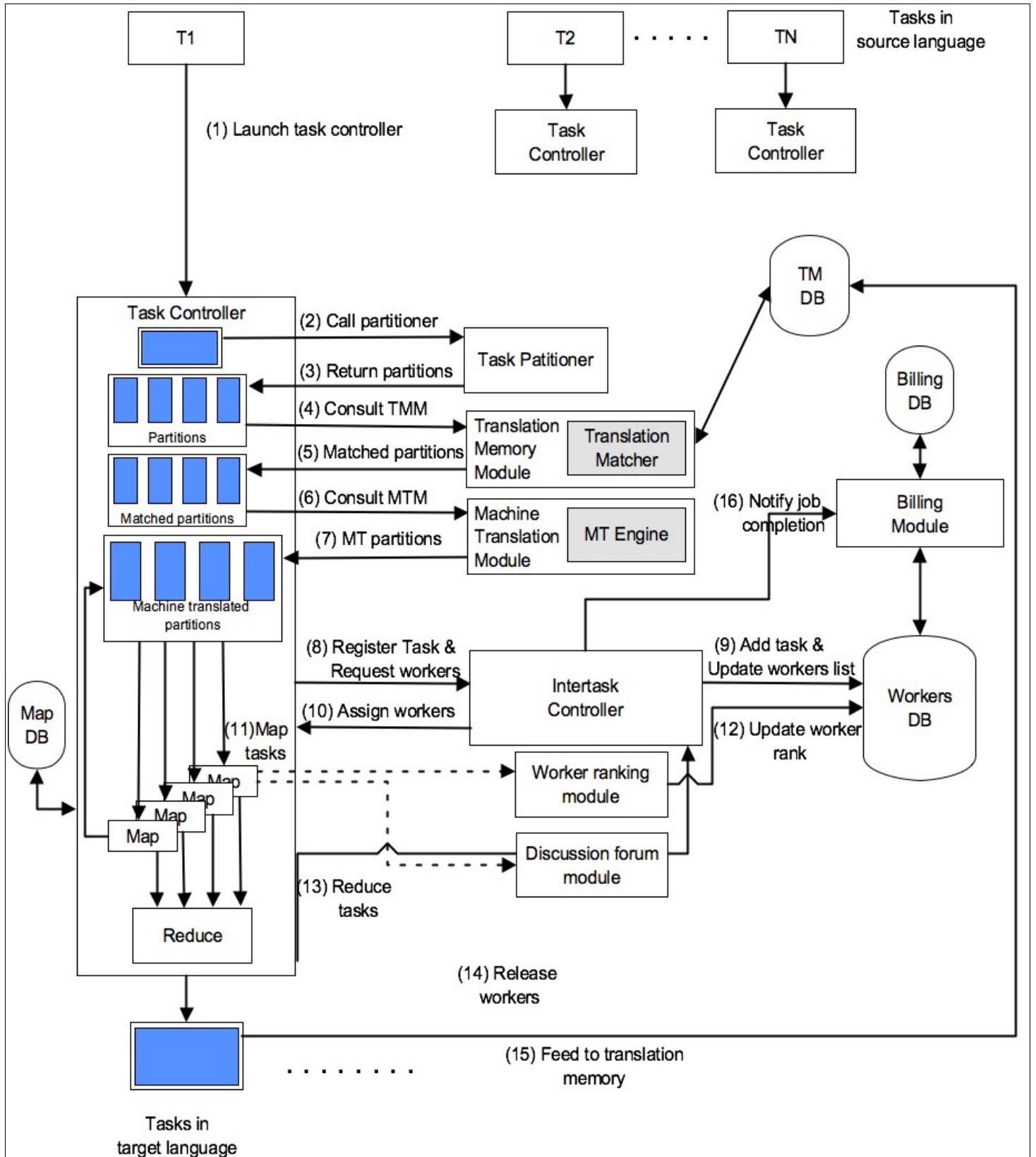
Some secondary considerations, that are not inserted in the current version of the architecture but may result in useful features, such in example:

Back up pool:

ITC can keep a back-up pool of workers for each task offer, where to save workers that are not assigned to any task but they are available and suit the requisites, so to have a quickly back-up person in case of resignation of some other workers.

Chapter 8 lists many suggestions and considerations which, for some reasons, are not present in the current design, but which may improve the designed system's performances.

## Overview of the Final System Architecture



## 4.5 Ranking System

What motivate people to work in a system as crowd sourcing? According to some interesting research done on crowd sourcing systems and worker's motivations [10] and [11] there are many reasons behind users participation, and those reasons differ accordingly to many criteria.

The designed architecture proposes a novel ranking methodology for crowd sourcing system that has been thought to motivate people to deliver good quality work by offering an increased amount of money which totally depends on user ranking.

The idea is that worker can increase his ranking by providing good quality work, and reach a certain level of ranking which can be monetized into a significant money income. Moreover, the ranking system is increasing the performance of the system. For example, if all the workers obtain such a high ranking that really few phases are needed to obtain quality in translation, this would speed up the velocity at which text are translated, and potentially save money.

The ranking of each worker is influenced by any task he performs in the system, but it has been also proposed to offer a series of test and certification to allow professionals to increase their ranking quickly. If it may exists something as a "CA Technologies Translator Certification", the objective of obtaining this certification may motivate professional translators to participate in the crowd sourcing system, potentially making the inscription to this system viral.

## 4.6 Action-Verification Unit

All the analysis and study done for this project has been used to redact the paper [25]. This paper start by discussing about criteria and characteristics that make some tasks more or less suitable to be solved with a crowd sourcing methodology. Some of those characteristics are the following:

- Activities must not be easy to automate: crowd sourcing activities usually require to perform actions that are better solved by the human brain than by currently available algorithms. Usual examples are tasks where creativity is essential, such as the proposal of new designs;
- Crowd sourcing offers intrinsic advantages: tasks where the geographic distribution of the individuals provides a higher quality access to information; or tasks where the complexity of the problems posed are so high that there are not predefined mechanisms to solve them. For example cultural barriers and localization process.
- Information involved in the process must not be confidential: since data must be sent to the crowd, processes that involve sensitive information are not suitable in general for this type of solution. On the positive side, crowd sourcing also mitigates concerns about loss of privacy, since a single provider does not have a global view of anyone's data.
- Training must be simple or highly automatic: complex training processes are not suitable for crowd sourcing since this would imply training thousands or even millions of people, which would be unaffordable.

Even if there are examples where crowd sourcing resulted the best possible option, crowd sourcing is often associated to low quality work. This may be because it is difficult to establish automatic mechanisms to evaluate work in a crowd-based systems, since, for many task suitable for crowd sourcing, the concept of quality is nuanced, as it may be as example creativity.

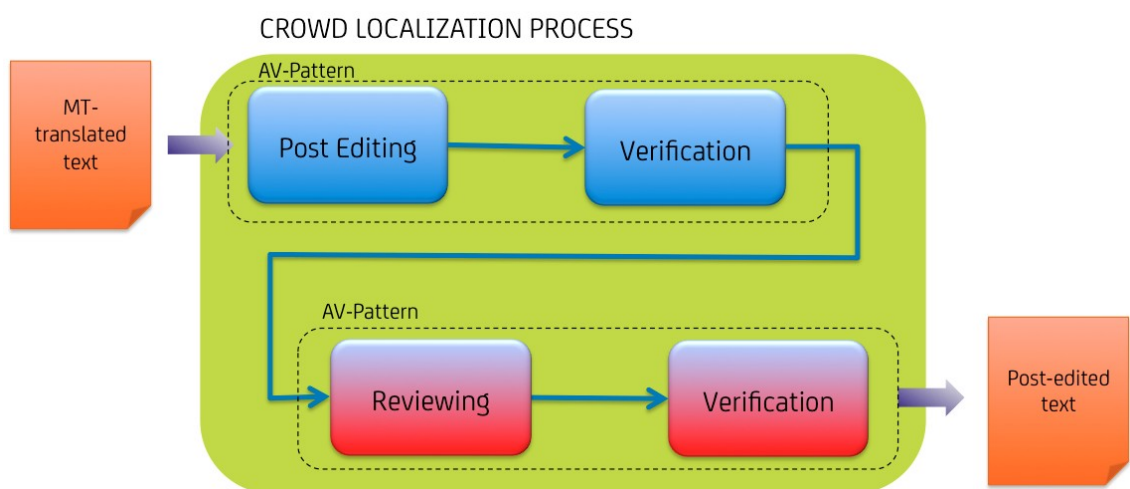
At the same time, humans beings are often the best judge and can evaluate the

quality of a work much better than any automatic system, especially when there is not any accepted quality measure for a specific work.

In this paper the methodology created during the project and described in the previous chapters have been formalized as an **Action Verification Unit**.

Therefore, the paper proposes an Action Verification Unit which is used to establishes a relationship pattern between workers of the crowd, and it helps them to work collaboratively and provide a higher degree of quality.

In general, in the Action phase a single worker perform some kind of work which is verified from a set of worker at the Verification phase. The average of the judgment provided at Verification phase is than compared with a threshold in order to decide if the worker at the Action phase has to redo the work done or not. Specifically for the Translation process described in the previous chapter, the system is composed by these AV-Unit



## 5 Prototyping

*"A prototype is a rudimentary working model of a product or information system, usually built for demonstration purposes or as part of the development process. In the systems development life cycle (SDLC) Prototyping Model, a basic version of the system is built, tested, and then reworked as necessary until an acceptable prototype is finally achieved from which the complete system or product can now be developed."*

According to waterfall SW engineering process, after prerequisites analysis and design of the architecture, it comes implementation.

In addition to comply with MTI master thesis requirements, a prototype of the described system has been implemented to perform some experiment and to study how different skilled workers influence the quality of the translation within the designed process.

So, more than a mere facade of the final system, the prototype has been created and used as a tool to study crowd environment and its behavior. More specifically, the prototype has been developed starting from the designed architecture, but it was decided to be a lighter version, less complex than what has been designed, but more flexible and manageable.

The main idea behind the creation of the prototype was to evaluate the designed process and the Action-Verification pattern, and so the prototype allows to assign differently skilled users to different translation steps described in the previous chapters, and it is more manageable in order to investigate the following issues:

-which tasks are more important than the others in providing good quality;

-At which step better users "influence" more the final translation, so to understand which task they should be assigned to;

-which combinations of users provide good quality;

-If quality in translation can be obtained just from skilled users, or even it can be provided by non-skilled users, or a combination of both.

The main idea of running experiment on a prototype is to check the potentials of the designed quality assurance method, using a reduced crowd, and to understand how the designed process of Action-Verification can be improved. In addition, it was decided to provide CA Technologies with tangible evidence that Crowd translation can reach CA Technologies quality standards of translation.

Mainly, the prototype allows the following tasks:

- User registration;
- Upload of text in «.ttx» format and parse of it into a Database;
- Management of the different translations steps;
- Management of task-user assignation;
- Inform users by email when they have been assigned to a task;
- Online Post edit phase;
- Online Verification phase;
- Online Review phase;

The coding of the prototype started in middle April 2012 and lasted until end of June 2012. During this period Java programming language has been used to code the prototype.

The creation of the prototype has been done with the collaboration of another student from UPC, and under the supervision of Victor Muntés, head of CA Labs Europe. Project management, organization and scheduling of: tasks, deadlines and meetings are also issues which have been addressed during the development of this project. Knowledge and facts learned during courses such as PESBD have been extensively applied to the organization of roles and scheduling of the different project milestone. More specifically, the main part of

the prototype I was in charge to develop is the creation of the interfaces used to work on the different phases of the translation process and the logic behind them, creating more than 50 files among different Servlets, jsp and java classes.

Other session of collaborative coding were done with the other student, especially for parts of the prototype where my code and his code needed to interact and work together. After an intensive coding period, never ending sessions of testing, bugs corrections, and retesting had taken few weeks more. During the evolution of the experiment, technical support for users and code adjustment have been also time consuming. The coding of the prototype required almost 3 months of working, and the most important issues of this work are explained later on in this chapter.

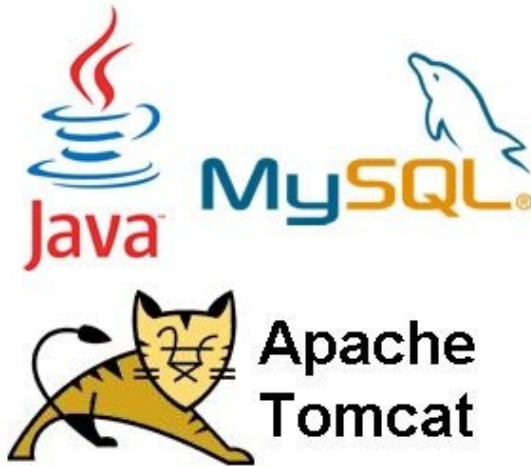
All the prototype is localized in Spanish and Catalan. The prototype is also localized in two more languages which are Chinese and German because, originally, experiment was thought to include those two language also, but the difficulty to reach native speakers and their scarcity during the registration period, made us to desist.

The prototype loads different jsp according to user preferences, meaning that users can choose a system totally localized in Catalan or Spanish, among the other languages. Since the prototype has been used for an experiment about Translation and Localization, it has been done an extra effort to present to user neat and grammatically correct interfaces and instructions. This process show one time more the Localization and Translation of a SW product is a continuous process fraught with difficulties.

This chapter explains some detail of the implementation, among with a panoramic of the prototype.

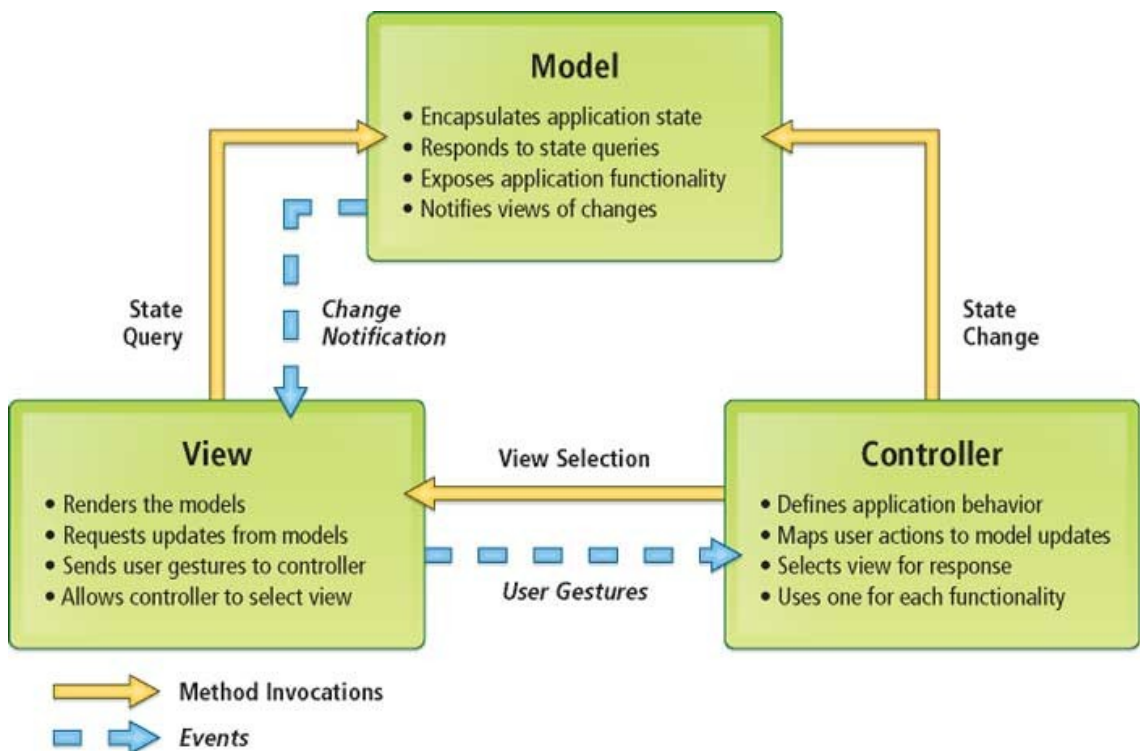


## 5.1 Technology



Since the prototype, basically, is a web application, it was decided to use Java programming language, which is the most used in web application especially due to its efficiency. Mysql technology is used as database, while the web part is running on an Apache Tomact web server.

Prototype has been coded following the Model – View - Controller architectural pattern, mainly because it has been studied in PROP course and also because it is one of the most popular programming pattern.



The Model is used to create the logic of the application and it is responsible to store data in memory and interact with the DataBase. View module is in charge of the output representation of the data obtained from the Model module, while the Controller module is the one which determines the application flow, handling events and interacting with the Model module to change attributes and, in general, the state of the system.

The computer used to deploy the prototype and running the experiment was provided by the Computer Architecture department of FIB (UPC), and it had the following HW characteristics:

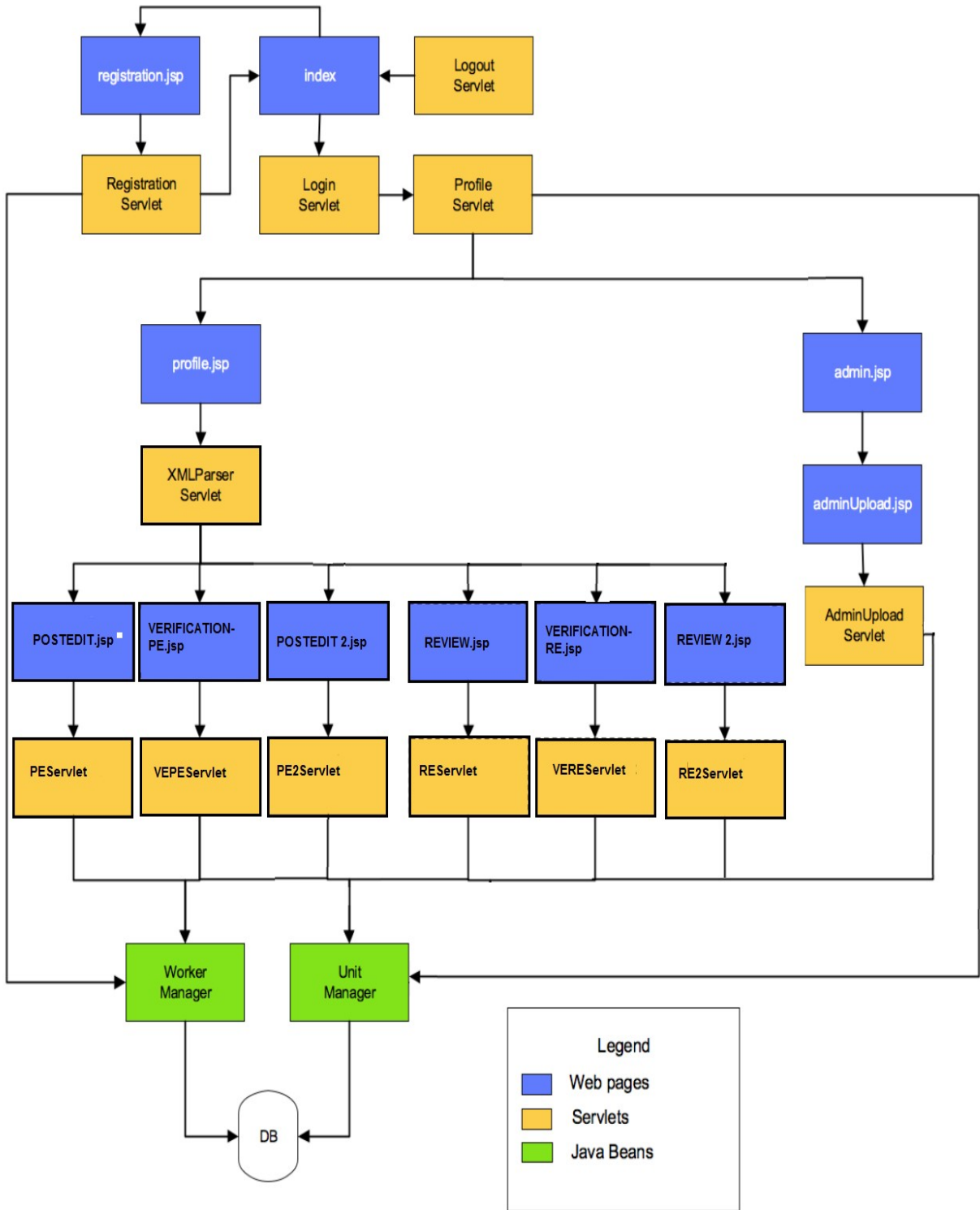
CPU Intel(R) Xeon(R) CPU E5606 2.13GHz;

10 GB RAM;

1 TB Hard disk;


OS Linux 3.0.0-17-Ubuntu Server 4.6.1;

## 5.2 Overview of the application's web flow



Prototype is currently reachable at URL: <http://crowd.pc.ac.upc.edu> . The web site presents the experiment, the reasons and motivations behind it. Web site also was used to promote the experiment by advertising that participants of the experiment have the possibility to win an Ipad.

# CrowdTranslate



## Participate and win an iPad2!

### Introduction

More than 6000 languages are in use worldwide, and while 1.8 billion people speak English, the other 5.2 billion do not. Translating products and services into many local languages is very important for business and educational use. However, due to the expense and time involved, delivering technologies in many geographies has been difficult.

Universitat Politecnica de Catalunya (UPC), through the DAMA-UPC research group and in collaboration with CA Labs, is involved in an exciting new experimental project, and wants to enlist the power of the crowd (you!) to help. In this project, we will combine the power of machines with the human capacity for language to provide high quality translations from English to other languages.

Are you interested in participating in this new technology? You do not need to know English nor have any previous experience in translation. Your role in the project will depend upon your profile, experience translating and your language skills. We need a lot of people! [Register in the system and we will contact you very soon.](#)

To encourage your participation, we will be giving away an iPad to one winner among the participants of the project.

### Sign in

Email

Password

[Login](#)

[New User?](#) [Register](#)

Users could register in the system via web.

## Formulari de registre

Nom \*

Any de naixement\*

Adreça electrònica\*

Contrasenya \*

Confirmar Contrasenya\*

---

Quina és la teva llengua materna?\*

Català  Castellà  Alemany  Xinès simplificat

Quin és el teu nivell de comprensió d'anglès? \*

Quina consideres que és la teva àrea d'expertesa principal? \*

Relacionada amb llengües: Literatura, Filologia, etc  
 Relacionada amb les TIC: Informàtica, xarxes, comunicació, etc  
 Relacionada amb les TIC i amb la traducció  
 Relacionada amb altres estudis universitaris  
 Altres

Quina experiència tens com a traductor? \*

Traductor professional  
 Estudiant de traducció  
 Cap

Creus que ets capaç d'entendre una guia d'usuari d'un programari (programes informàtics amb la documentació corresponent)? \*

Sí  
 No

Prement el botó "Vull participar a l'experiment", estàs acceptant la [Política de Privacitat](#).

After completing the registration steps, user could connect to his profile page and see which task he was assigned to.

UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH

in collaboration with

ca CA Labs

funded by

FEDER GOBERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD INVESTIN SPAIN

INICI PERFIL SORTEIG DE L'IPAD FAQ SORTIR

**Tasques Pendents**

Tens les següent tasques pendents. Fes clic a "Acceptar" per començar a treballar-hi.

Id	Estat	
ct1-224-308	pending	Acceptar

COPYRIGHT (C) 2012 WWW.CROWD.PC.AC.UPC.EDU. ALL RIGHTS RESERVED. IF YOU ENCOUNTER ANY PROBLEMS DURING THE EXPERIMENT, CONTACT US AT SUPORT.DAMA@AC.UPC.EDU

An administration account was created to control the evolution of the experiment. The administration account has a different profile page where different Servlets are used to manage the experiment. With the administration control panel it is possible to assign tasks to users.

Dashboard

PE Jobs PEVE Jobs RE Jobs REVE Jobs

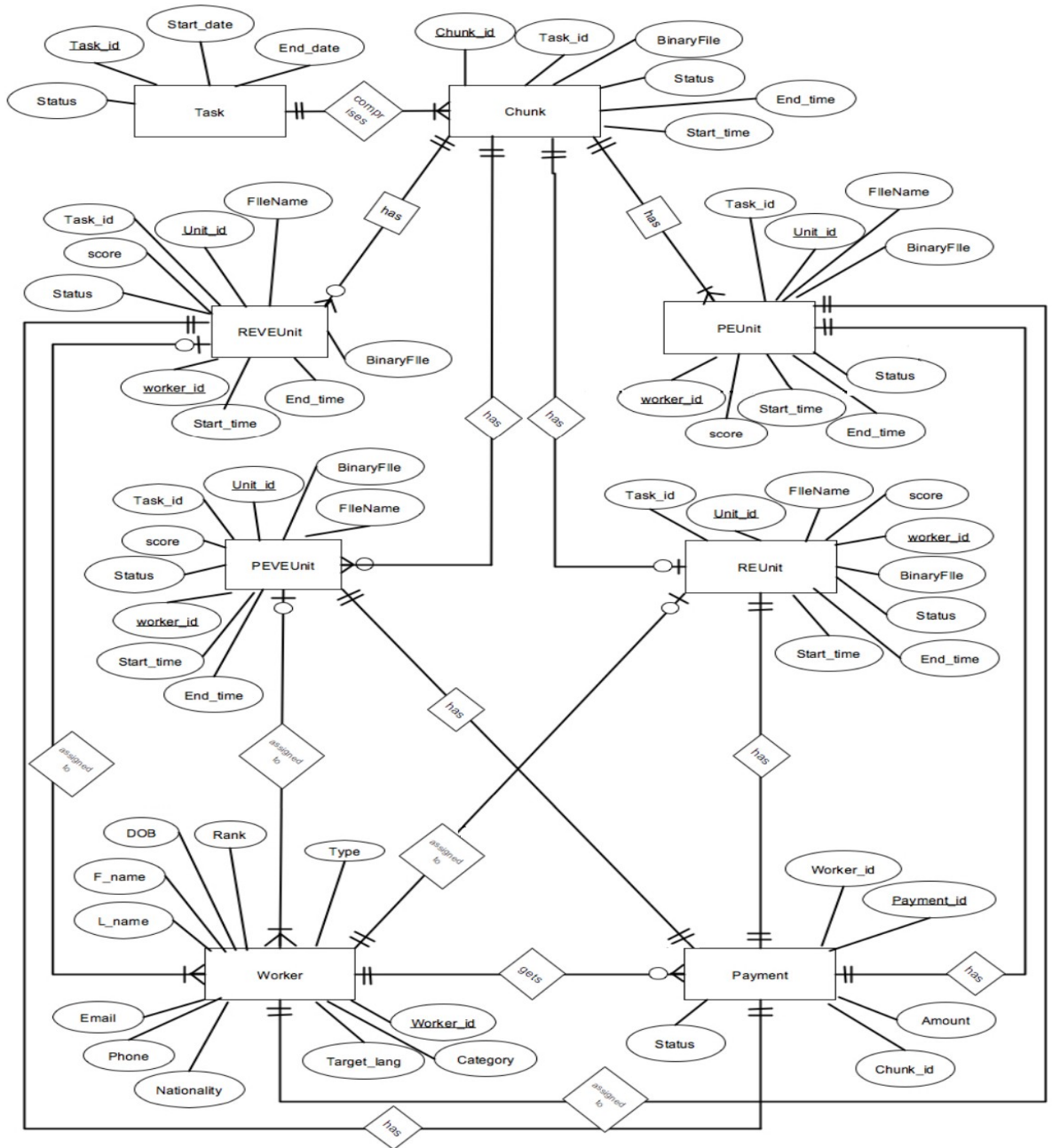
**Post Editing Tasks**

Unit Id	Worker id	Start time	End time	Status	Revised Status
ct1-152	152	2012-06-23 03:17:13	2012-06-23 10:30:48	verifying	none
ct1-155	155	2012-06-23 03:17:16	0000-00-00 00:00:00	pending	none
ct1-165	165	2012-06-23 03:17:16	0000-00-00 00:00:00	pending	none
ct1-224	224	2012-06-23 03:17:14	0000-00-00 00:00:00	pending	none
ct1-252	252	2012-06-23 03:17:11	2012-07-05 13:39:14	verifying	none
ct1-290	290	2012-06-23 03:17:12	2012-07-02 17:38:24	verifying	none
ct1-305	305	2012-06-23 03:17:17	0000-00-00 00:00:00	pending	none
ct1-319	319	2012-06-23 03:17:17	2012-07-03 15:16:22	verifying	none
ct1-340	340	2012-06-23 03:17:12	2012-07-12 12:18:05	verifying	none
ct1-349	349	2012-06-23 03:17:14	0000-00-00 00:00:00	pending	none
ct1-378	378	2012-06-23 03:17:13	2012-06-29 16:15:06	verifying	none
ct1-437	437	2012-06-23 03:17:15	2012-07-06 12:08:52	verifying	none
es1-247	247	2012-06-23 03:15:55	0000-00-00 00:00:00	pending	none
es1-256	256	2012-06-23 03:15:58	0000-00-00 00:00:00	pending	none
es1-269	269	2012-06-23 03:15:59	0000-00-00 00:00:00	pending	none
es1-276	276	2012-06-23 03:16:00	0000-00-00 00:00:00	pending	none
es1-300	300	2012-06-23 03:16:01	0000-00-00 00:00:00	pending	none
es1-348	348	2012-06-23 03:16:01	2012-06-28 22:04:16	waiting	Allocate PEVE
es1-370	370	2012-06-23 03:15:56	2012-07-06 14:20:11	waiting	Allocate PEVE
es1-375	375	2012-06-23 03:15:57	2012-06-26 21:23:55	waiting	Allocate PEVE
es1-376	376	2012-06-23 03:15:58	2012-07-06 13:16:28	waiting	Allocate PEVE
es1-388	388	2012-06-23 03:15:59	0000-00-00 00:00:00	pending	none
es1-466	466	2012-06-23 03:16:00	0000-00-00 00:00:00	pending	none
es1-469	469	2012-06-23 03:16:02	0000-00-00 00:00:00	pending	none

When a task is assigned to user, the mailer module sends him an email via the UPC's mail server *rely.upc.edu* , informing user that he has been assigned to a task. The mailer sends different mail in Spanish or Catalan according to user preferences.

### 5.3 Data Base

The scope of the prototype was not entirely defined since the beginning of the coding period, and for this reason the database schema used for the prototype has had more attributes than what it really needed.



(Entity Relations diagram of the tables used to manage the experiment)



Unit tables such, in example, PEUnit, VEPEUnit etc, contain information on user-task assignment.

Each user-task assignation is identified with the following Unit\_id schema:

TargetLanguage – Worker\_id ( – Worker\_id )

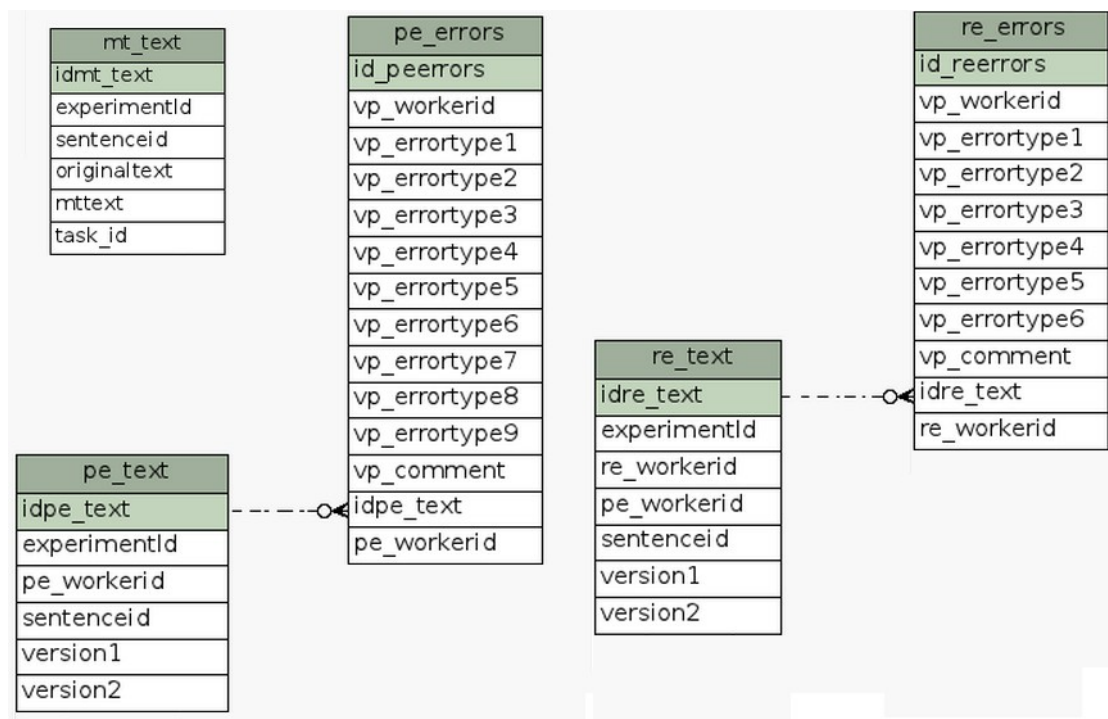
For example, to assign worker number 2 to a Verification of PE phase to verify the text produced by worker number 1, within a Catalan experiment, the Unit\_id would be the following: ct-1-2 .

PEUnit	PEVEUnit	REUnit	REVEUnit	worker
unit_id	unit_id	unit_id	unit_id	worker_id
task_id	task_id	task_id	task_id	f_name
worker_id	worker_id	worker_id	worker_id	dob
binaryfile	binaryfile	binaryfile	binaryfile	email
filename	filename	filename	filename	password
start_time	start_time	start_time	start_time	target_lang
end_time	end_time	end_time	end_time	english_level
status	status	status	status	expertise
score	score	score	score	rank
vbinaryfile	pe_id	vbinaryfile	re_id	experience
vstatus		vstatus		tech
re_visibility		re_visibility		category
pe_visibility		vstart_time		type
vstart_time		vend_time		status
vend_time				grp
				gr
				es
				cat
				ch
				keep_data
				multi_lang
				allow_login
				test_completed
				number_of_tasks

The data user provides at registration time are saved in the table worker, while the answers given in the initial test and user final score were saved in the initial\_test table.

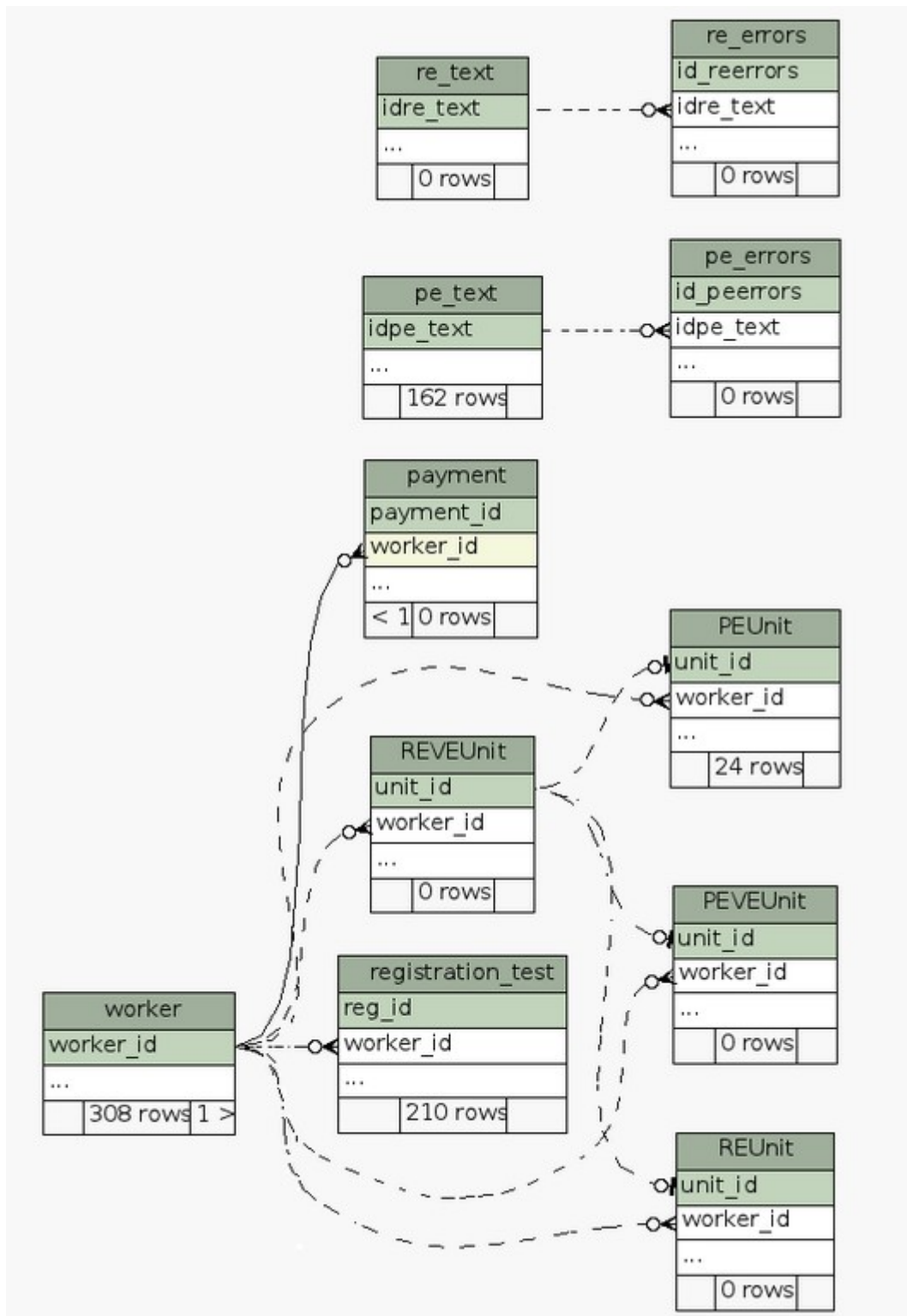
Tables MT\_text, PE\_text and RE\_text contain correspondingly the machine translation text, the post edited text and the reviewed text, while PE\_errors and RE\_errors contain the errors found during the correspondent verification phases.

The different texts, errors, comments and, in general, all the material produced by users is saved in the following tables:



As it can be seen, the type of errors users can find at Verification phases are limited to 9 in the Verification of Post edit; and to 6 in the Verification of Review. It was decided not to give an open set of possible errors to Verificators so to have common errors found by different Verificator. At the same time, each Verificator could add a comment to each sentence he found not to be correct, and those comments are shown to the other users during phase Post edit 2 or Review 2.

## Overview of the Database



## 5.4 Parser

The development of the prototype started with the creation of a parser. During the prerequisite analysis, it was decided the system would be able to interact with technology used by CA Technologies Localization department, specifically its translation tool, Trados SDL translation tool\*.

Trados SDL is one of the most used software in the Translation and Localization world, and its files are formatted with a proprietary format similar to an XML file where each file contains the original text and the machine translated text. Each sentence is contained between the tokens <TU> and </TU>, which contains internally other tokens used to keep attributes of the sentence, such, in example, the token <TUV Lang="EN-US"> which indicates that the text contained in that token is text written in American English.

```
<TRADOStag Version="2.0">
- <FrontMatter>
  <ToolSettings CreationToolVersion="1.0" CreationTool="SDLX XConvert" CreationDate="20120312T033116Z"/>
  <UserSettings SourceDocumentPath="C:\Documents and Settings\cusil01\Desktop\ilaria\centr apps\TASK34797_EN-US-IT_CA_it_IT_2746_20120223150236_ASEdge--Protection Manager Application-HTML_PDF-PM_xml.itd" TargetDefaultFont="Arial" TargetLanguage="IT-IT" SourceLanguage="EN-US" SettingsPath="C:\Program Files\SDL International\SDLX\opentag.ini" SettingsName="OpenTag" O-Encoding="windows-1252" DataType="XML"/>
</FrontMatter>
- <Body>
  - <Raw>
    <ut DisplayText="HTML" RightEdge="angle" Style="external" Type="start"><HTML></ut>
    <ut DisplayText="HEAD" RightEdge="angle" Style="external" Type="start"><HEAD></ut>
  - <Tu>
    - <Tuv Lang="EN-US">
      <ut DisplayText="META" >META http-equiv=Content-Type content="text/html; charset=windows-1252"></ut>
    </Tuv>
    - <Tuv Lang="IT-IT">
      <ut DisplayText="META" >META http-equiv=Content-Type content="text/html; charset=windows-1252"></ut>
    </Tuv>
  </Tu>
  <ut DisplayText="HEAD" Style="external" Type="end" LeftEdge="angle"></HEAD></ut>
  <ut DisplayText="BODY" RightEdge="angle" Style="external" Type="start"><BODY><table cellspacing="1" style="border-collapse: collapse" bordercolor="#111111" border="1" cellpadding="2" width="100%"></ut>
  <ut DisplayText="P id="12" RightEdge="angle" Style="external" Type="start"><tr><td width="10%" bgcolor="#C0C0C0" valign="top"><p align="right">12</td><td width="90%"></ut>
- <Tu>
  <Tuv Lang="EN-US">How to Uninstall CA ARCserve Central Protection Manager</Tuv>
  <Tuv Lang="IT-IT">How to Uninstall CA ARCserve Central Protection Manager</Tuv>
</Tu>
  <ut DisplayText="P" Style="external" Type="end" LeftEdge="angle"></td></tr></ut>
  <ut DisplayText="P id="13" RightEdge="angle" Style="external" Type="start"><tr><td width="10%" bgcolor="#C0C0C0" valign="top"><p align="right">13</td><td width="90%"></ut>
- <Tu>
  <Tuv Lang="EN-US">How to Uninstall CA ARCserve Central Protection Manager</Tuv>
  <Tuv Lang="IT-IT">How to Uninstall CA ARCserve Central Protection Manager</Tuv>
</Tu>
  <ut DisplayText="P" Style="external" Type="end" LeftEdge="angle"></td></tr></ut>
  <ut DisplayText="P id="14" RightEdge="angle" Style="external" Type="start"><tr><td width="10%" bgcolor="#C0C0C0" valign="top"><p align="right">14</td><td width="90%"></ut>
```

\*[[www.trados.com](http://www.trados.com)]

The Parser written for extracting the text from this particular format uses a class called: “DocumentBuilderFactory” that enables the application to obtain a DOM object tree\*, which then is navigated with function “Xpath”,\*\* looking for tokens that contain text. Since some tokens contained nested tokens, especially when one sentence in the original language corresponded to more than one sentences in the target language, function *getAllChildNodes* is used to traverse recursively the DOM tree.

The parser class is able to parse those specific files and extract the original language and the machine translated text into two separated files. This class is used by Upload Servlet, which is reachable from the administration control panel of the prototype.

### Administrator Page.

Parse file for a new experiment

Upload an XML:  No se ha seleccionado ningun archivo

View Dashboard

The Upload Servlet loads the selected file, parses it, separates the original text and the machine translated text and then saves those two distinct texts in the MT\_text table of the database.

\*[<http://docs.oracle.com/javase/1.4.2/docs/api/javax/xml/parsers/DocumentBuilderFactory.html> ]

\*\*[[http://www.w3schools.com/xpath/xpath\\_functions.asp](http://www.w3schools.com/xpath/xpath_functions.asp) ]

## Table crowd.mt\_text

Related columns  Constraints  Comments  Legend

Column	Type	Size	Nulls	Auto	Default	Children	Parents	Comments
idmt_text	int	11		√				
experimentId	varchar	45						
sentenceid	varchar	45						
originaltext	text	65535						
mttext	text	65535						
task_id	tinyint	4	√		null			


Table contained 54 rows at Tue Jun 26 20:07 CEST 2012

### Indexes:

Column(s)	Type	Sort	Constraint Name
idmt_text	Primary key	Asc	PRIMARY


When the text has been uploaded, an experiment is created and users can be assigned to the first phase of the designed translation process: Post Edit of a text.

When user logs in the system via the web page, Profile Servlet checks worker status in the worker table by taking the worker ID which is passed as a session variable. If user's status is set as busy, it means that user has been assigned to a task. So Profile Servlet checks worker attribute "type" (PE,VEPE,...) and then query the corresponding task table (PEUnit , PEVEUnit...). When a task assigned to user is found in the univ table, it is shown on user personal page, and user can click the button to start the task.







**UNIVERSITAT POLITÈCNICA DE CATALUNYA**  
BARCELONATECH

in collaboration with



**CA Labs**

funded by

INICI
PERFIL
SORTEIG DE L'IPAD
FAQ
SORTIR

**Tasques Pendents**

Tens les següent tasques pendents. Fes clic a "Acceptar" per començar a treballar-hi.

Id	Estat	
ct1-224-308	pending	<span style="background-color: #0070C0; color: white; padding: 5px 10px; border-radius: 5px; font-weight: bold;">Acceptar</span>

COPYRIGHT (C) 2012 WWW.CROWD.PC.AC.UPC.EDU. ALL RIGHTS RESERVED.  
IF YOU ENCOUNTER ANY PROBLEMS DURING THE EXPERIMENT, CONTACT US AT SUPPORT.DAMA@AC.UPC.EDU

## **5.5 Post Edit phase**

As explained in the previous chapter, Post edit phase is the first step of the translation process. At this step, one worker from the crowd is assigned to this phase where he has to work on the machine translated text and improve it. The interface was designed resembling Trados SDL tool, where two columns are displayed, and the machine translated text is contained in the left column while the original text is in the right column.

Even if this may seem untoward, professional translators from CA Technologies suggested that having the original text on the right part of the screen instead of having it on the left, allow users to perform a better translation and to be less influenced from the original text.

Therefore, user has to modify the text contained in the left column and, whenever he feels like, he can save or submit the work. If user press the save button, he can log out and, at any time he logs in again, he can continue to work from the same point he saved. Analogously, user can submit the text whenever he consider he has improved it enough, and he feels text can not be improved anymore.

# CrowdTranslate

## Fase de postedició

En aquesta pàgina hi ha dues columnes: la de l'esquerra mostra el text traduït automàticament i la de la dreta mostra el text original en anglès que s'ha de traduir al català. L'objectiu és millorar (si cal) el text de l'esquerra per expressar correctament el sentit del text original.

**Siusplau, consulta la guia d'estil per fer aquesta tasca.**

-  [Style guide.pdf](#)
-  [Glossary.pdf](#)
-  [Original file.pdf](#)

Text traduït amb traducció automàtica per corregir i millorar

Text original

Instal·lació de la Passarel·la de mesura (MGT) i Configuració	Metering Gateway Installation and Configuration
L'aplicació MGT és una passarel·la que metering que accepta dades de mesura des de grids dins del datacenter d'un client i envia aquelles dades al sistema de mesura	The MGT application is a metering gateway that accepts metering data from grids within a customer's datacenter and forwards that data to <productname>'s metering
Verifiqui que cap missatge d'error no es registra al quadre de comandaments de grid.	Verify no error messages are logged to the grid dashboard.
Si hi ha missatges d'error, sisplau contacti suport tècnic.	If there are error messages, please contact technical support.
Copyright © 2011.	Copyright © 2011.
Tots els drets reservats.	All rights reserved.

Guardar

Guardar i Enviar



When PEDetailServlet is called, it firstly performs different queries in PE\_text table to check if user has done some precedent work or if it is the first time he starts the task: if user has already some text saved in the table, it means user has already saved text and so PEDetailServlet loads user previous text and allows user to continue his precedent work. Else way, if it is the first time user has started the task, PEDetailServlet queries MT\_text table looking for the text to which experiment user has been assigned, and save it in the PE\_text table, within user worker\_id. Consequently it loads the same text and it passes it over to the View module which shows it in the interface.

```
String q = "SELECT `pe_text`.`idpe_text`, `pe_text`.`experimentId`, `pe_text`.`sentenceid`, `mt_text`.`originaltext`, "  
+ "`mt_text`.`mttext`, `pe_text`.`version1` FROM `mt_text` "  
+ "left join `pe_text` "  
+ "on `pe_text`.`experimentId` = `mt_text`.`experimentId` "  
+ "and `pe_text`.`sentenceid` = `mt_text`.`sentenceid` "  
+ "where `mt_text`.`experimentId` = '"  
+ experimentid  
+ "'" "  
+ " and `pe_text`.`pe_workerid` = '" + workerid + "'";  
PETranslationManager ptrm = new PETranslationManager(q);
```

The View module is composed by different jsp, and PEDetailServlet calls PEDetail.jsp to create the main structure of the interface.

When user completes the task by pressing the submit button, the PEDetailServlet simply does an updated in the PE\_text table and the task is removed from worker's actual queue, by updating the status in the PEUnit table. Worker at Post edit phase is not set as 'free' yet but instead is set as "waiting", because he may have to modify the text during phase Post edit 2.

Whenever user presses the submit button, the task is considered completed, so task is removed from user's queue and the text is saved as in the PE\_text table, column: "version 1".

## Table crowd.pe\_text

Implied relationships  Related columns  Constraints  Comments  Legend

Column	Type	Size	Nulls	Auto	Default	Children	Parents
idpe_text	int	11		√		<a href="#">pe_errors.idpe_text</a>	
experimentId	varchar	45					
pe_workerid	int	11					
sentenceid	int	11					
version1	text	65535	√		null		
version2	text	65535	√		null		

Table contained 162 rows at Tue Jun 26 20:07 CEST 2012

### Indexes:

Column(s)	Type	Sort
idpe_text	Primary key	Asc

At this point, a user has improved the machine translated text, which is ready to be verified by a group of Verificators during the Verification of Post Edit .

## 5.6 Verification of Post edit phase

The output from the Post edit phase is passed to phase Verification of Post Edit, where multiple workers verify the text produced from a worker at Post edit phase. Workers at Verification phase do not have to correct errors nor modify the text in any way, but they just have to find errors in the sentence and select the appropriate check box. To underline errors and provide their suggestion, users may leave comments to support and explain their verification.

This methodology follows the Action-Verification pattern explained in previous chapter, where the task performed by someone is subjected to a control group. Similarly to Post edit phase, the interface shows two columns with the original text contained in the left column, and the post edited text contained in the right column.

While original text is used by the Verificator to check errors in translation such, for example, the meaning of the sentence was not modified by the Post Editor, the post edited text is revised by the Verificators which look for grammar, typos and other errors. Each Verificator can be assigned to one or more tasks to verify various Post Edited texts provided by different users.

Since this task requires less time to be completed, and supposedly it is also easier, worker can not save the work, and he can just submit it. When user submits the task, errors and comments are inserted in the PE\_errors table.

# CrowdTranslate

## Revisió

Pren-te el teu temps per tal de contestar a les errors de frases que es plantegen en aquesta pàgina. Una vegada facis clic al botó "Emmagatzema la resposta i continua" ja que no podràs modificar les teves respostes i el sistema emmagatzemarà els resultats a la nostra base de dades. Durant la prova, pots consultar la guia d'estil o qualsevol altra font d'informació. Quant més investiguis millor qualitat del teu treball!

Per aquest test no hi ha cap llistat de termes disponible. Si tens algun dubte consulteu els enllaços que us proveïm a la guia d'estil.

**ATENCIÓ:** és molt important que segueixis les indicacions de la guia d'estil. Et recomanem que imprimeixis el document. El pots fer servir tant durant la prova com més tard durant l'experiment.

- Guia d'estil.pdf
- Glossary.pdf
- File Original.pdf

Frase original a traducir

Texto traducido a evaluar

Metering Gateway Installation and Configuration

Instal·lació i Configuració de la Passarel·la de mesura (MGT)

- Opcions de software
  Omissió/Addició
  Error de traducció
  Error ortogràfic
  Puntuació
  Estil
  Gramàtica
  Terminologia
  Text no traduïble

Deixa el teu comentari aquí

The MGT application is a metering gateway that accepts metering data from grids within a customer's datacenter and forwards that data to <productname>'s metering system.

L'aplicació MGT és una passarel·la de mesura que accepta dades de mesura provinents de grids dins del centre de dades d'un client i les envia al sistema de mesura de <productname>.

- Opcions de software
  Omissió/Addició
  Error de traducció
  Error ortogràfic
  Puntuació
  Estil
  Gramàtica
  Terminologia
  Text no traduïble

Deixa el teu comentari aquí

All rights reserved.

Tots els drets reservats.

- Opcions de software
  Omissió/Addició
  Error de traducció
  Error ortogràfic
  Puntuació
  Estil
  Gramàtica
  Terminologia
  Text no traduïble

No s'ha seguit la terminologia indicada al glossari.

Deixa el teu comentari aquí

Guarda i Envia

When user logs in, the Profile Servlet performs the same set of actions as explained in the previous chapter and, if the user is assigned to Verification of Post Edit task, Profile Servlet calls VPDetailServlet which performs an inner query between MT\_text table and the PE\_text table, to load, respectively, the Original text and the Post Edited text from the correct user. VPDetailServlet knows which post edited text it has to be loaded by checking the Unit id which, as explained previously, is made correspondingly by the language experiment, the post editor worker ID and the Verificator worker ID, as example: ct-1-2 .

```
String q2 = "SELECT `pe_text`.`idpe_text`, `pe_text`.`sentenceid`, `mt_text`.`originaltext`, "
+ "`pe_text`.`version1` "
+ "FROM `pe_text` "
+ "inner join `mt_text` "
+ "on `pe_text`.`sentenceid` = `mt_text`.`sentenceid` "
+ "and `pe_text`.`experimentId` = `mt_text`.`experimentId` "
+ " where `pe_text`.`pe_workerid` = '" + pe_workerid + "' ";
ptrm.execute(q2);

Translation tr = null;
while (ptrm.moveToNext()) {
    tr = new Translation();
    tr.setId(ptrm.getPetranslationId());
    String original = null;
    try {
        original = ptrm.getResultSet().getString("originaltext");
    } catch (SQLException e) {
        e.printStackTrace();
    }
    tr.setDefaultLanguage(original);
    tr.setVersion1(ptrm.getVersion1());

    list.add(tr);
}
```

Then, VPdetail.jsp is called from the Servlet to create the structure of the interface shown in the previous image.

If error or comments have been found, those error are going to be shown to worker at Post edit phase, during the Post edit 2 phase.

When User submits the task, his status for the submitted task is changed from 'pending' to done in the PEVEUnit table.

When user feels he has found all the errors and submits the task by pressing the submit button, each error and comment is saved in the table PE\_errors, and each set of found errors and comments are related to the sentence they belong to with the foreign key “pe\_text.idpe\_text” .

### Table crowd.pe\_errors

Implied relationships  Related columns  Constraints  Comments  Legend

Column	Type	Size	Nulls	Auto	Default	Children	Parents
id_peerrors	int	11		√			
vp_workerid	int	11	√		0		
vp_errortype1	bit	1	√		b'0'		
vp_errortype2	bit	1	√		b'0'		
vp_errortype3	bit	1	√		b'0'		
vp_errortype4	bit	1	√		b'0'		
vp_errortype5	bit	1	√		b'0'		
vp_errortype6	bit	1	√		b'0'		
vp_errortype7	bit	1	√		b'0'		
vp_errortype8	bit	1	√		b'0'		
vp_errortype9	bit	1	√		b'0'		
vp_comment	text	65535	√		null		
idpe_text	int	11	√		null		<i>pe_text.idpe_text</i>
pe_workerid	int	11					

Table contained 0 rows at Tue Jun 26 20:07 CEST 2012

#### Indexes:

Column(s)	Type	Sort
id_peerrors	Primary key	Asc

At this point, some errors and comments may be left by Verificators, and if that is the case, those errors and comments are shown to the user whom produced the post edited text during the phase Post Edit 2.

When all the Verificators have completed their task and verified the text, phase Post Edit 2 can be started from the Administration control panel. During the experiment, 3 Verificators were assigned to verify 1 Post Edited text.

Dashboard

[PE Jobs](#)
[PEVE Jobs](#)
[RE Jobs](#)
[REVE Jobs](#)

PE Verification Jobs

Unit Id	Status	Unit Id	Status	Unit Id	Status	
ct1-152-175	done	ct1-152-203	done	ct1-152-452	pending	<a href="#">Send</a>
ct1-155-237	done	ct1-155-401	pending	ct1-155-436	pending	<a href="#">Send</a>
ct1-224-435	pending	ct1-224-308	done	ct1-224-209	done	<a href="#">Send</a>
ct1-252-175	done	ct1-252-203	done	ct1-252-452	pending	<a href="#">Send</a>
ct1-290-387	done	ct1-290-351	pending	ct1-290-188	done	<a href="#">Send</a>
ct1-319-157	done	ct1-319-233	pending	ct1-319-417	done	<a href="#">Send</a>
es1-326-318	pending	es1-326-459	pending	es1-326-271	pending	<a href="#">Send</a>
ct1-340-412	pending	ct1-340-147	done	ct1-340-451	pending	<a href="#">Send</a>
es1-348-453	pending	es1-348-359	pending	es1-348-297	pending	<a href="#">Send</a>
ct1-349-147	done	ct1-349-412	pending	ct1-349-451	pending	<a href="#">Send</a>
es1-370-271	pending	es1-370-459	pending	es1-370-318	pending	<a href="#">Send</a>
es1-375-251	done	es1-375-388	pending	es1-375-250	pending	<a href="#">Send</a>

## **5.7 Post Edit phase 2**

Errors found by Verificator in the previous phase are shown to worker during phase Post Edit 2, and worker assigned at Post Edit phase is asked to fix those errors and provide a final version of the text. The interface for the Post edit phase 2 shows only the sentences which contain errors or have some comments, so to make this task faster and easier for the worker. For each sentence, the different errors and comments from others users are shown all together in the same table.

Similarly at the first phase Post edit, worker at phase Post edit 2 can save or submit the work. The text is saved in the same table: PE\_text, but in the column "version 2". The two version of the same text (before verification and after verification) are kept in the database on purpose, to see how the verification phase influences the work done at Post edit phase. Worker at phase Post edit 2 can take into consideration the comments from the other workers, or if he does not agree with the reported errors, he can just submit the same text without any modification.



Text original
La teva traducció

To install and configure the MGT application, perform the following:

Per instal·lar i configurar l'aplicació MGT fes:

**Comentari de el crowd**

	Revisor	Error de traducció	Omissió/ Addició	Opcions de software	Gramàtica	Estil	Puntuació	Error ortogràfic	Text no traduïble	Terminologia	Comentari
203			X	X					X		Per instal·lar i configurar l'aplicació MGT, feu el següent:
175				X							canviaria el 'fes' per 'realitzi'

...

Basant-te en aquests comentaris, escriu la versió final de la traducció (deixa-ho buit si no fas cap canvi)

➔

---

Text original
La teva traducció

Download the application and copy it to the <productname>'s distro.

Descarrega l'aplicació i copia-la dins de la distribució de <productname>.

**Comentari de el crowd**

	Revisor	Error de traducció	Omissió/ Addició	Opcions de software	Gramàtica	Estil	Puntuació	Error ortogràfic	Text no traduïble	Terminologia	Comentari
203											Proposta: Descarregueu l'aplicació i copieu-la a la distribució
175											Correcte

...

Basant-te en aquests comentaris, escriu la versió final de la traducció (deixa-ho buit si no fas cap canvi)

➔

Guardar
Guardar i Enviar

When P2DetailServlet is called by Profile Servlet, it queries the PE\_errors table looking for all the errors found from the other worker on its text, and loads them, along with worker's post edited text from the PE\_text table and the original text from MT\_text table.

```

String q = "select distinct `pe_text`.`idpe_text`,`mt_text`.`originaltext`, `pe_text`.`version1`, `pe_text`.`version2`"
+ "from `mt_text`"
+ "left join `pe_text`"
+ "on `pe_text`.`experimentId` = `mt_text`.`experimentId`"
+ "and `pe_text`.`sentenceid` = `mt_text`.`sentenceid`"
+ "inner join `pe_errors`"
+ "on `pe_errors`.`idpe_text` = `pe_text`.`idpe_text`"
+ "where `pe_text`.`pe_workerid`=" + worker_id + " ";
PETranslationManager ptrm = new PETranslationManager(q);
Translation tr = null;
model.Error error = null;
ErrorManager em = new ErrorManager();

while (ptrm.moveToNext()) {
    tr = new Translation();
    tr.setId(ptrm.getPettranslationId());
    String original = null;
    try {
        original = ptrm.getResultSet().getString("originaltext");
    } catch (SQLException e) {
        e.printStackTrace();
    }
    tr.setDefaultLanguage(original);
    tr.setVersion1(ptrm.getVersion1());
    tr.setVersion2(ptrm.getVersion2());

    translationList.add(tr);

    q = "SELECT `pe_errors`.`id_peerrors`, "
+ "`pe_errors`.`vp_workerid`, "
+ "`pe_errors`.`vp_errortype1`, "
+ "`pe_errors`.`vp_errortype2`, "
+ "`pe_errors`.`vp_errortype3`, "
+ "`pe_errors`.`vp_errortype4`, "
+ "`pe_errors`.`vp_errortype5`, "
+ "`pe_errors`.`vp_errortype6`, "
+ "`pe_errors`.`vp_errortype7`, "
+ "`pe_errors`.`vp_errortype8`, "
+ "`pe_errors`.`vp_errortype9`, "
+ "`pe_errors`.`vp_comment`, "
+ "`pe_errors`.`idpe_text`, "
+ "`pe_errors`.`pe_workerid` " + "FROM `pe_errors` "
+ "where `pe_workerid`=" + worker_id + " "
+ "and `idpe_text` = " + tr.getId() + " ";
    em.execute(q);
    ArrayList<Error> tmpErrorList = new ArrayList<Error>();
    while (em.moveToNext()) {
        error = new Error();
        error.setId(em.getIderrors());
        error.setReviewId(em.getVpWorker());
        error.setErrorType1(em.getVpErrorType1());
        error.setErrorType2(em.getVpErrorType2());
        error.setErrorType3(em.getVpErrorType3());
        error.setErrorType4(em.getVpErrorType4());
        error.setErrorType5(em.getVpErrorType5());
        error.setErrorType6(em.getVpErrorType6());
        error.setErrorType7(em.getVpErrorType7());
        error.setErrorType8(em.getVpErrorType8());
        error.setErrorType9(em.getVpErrorType9());
        error.setComment(em.getVpComment());
        tmpErrorList.add(error);
    }
    errorList.put(String.valueOf(tr.getId()), tmpErrorList);
    em.closeConnection();
}
}

```

Then the P2Detail.jsp is called to create the structure of the page, and the jsp calls P2DetailTag class which creates the column and the table with the errors. After the worker has corrected the errors, the text is ready to continue to the following phase in the translation process: The review; done by a worker with its native language.

## 5.8 Review

From this phase on, monolingual users start to play. The text presented at this phase is the output text from the previous phases, and it is only written in the target language. User at Review phase is presented with a similar interface as in Post edit, where in the left column is contained the text he has to modify, while in the right column there is the same text mirrored, that can be used as reference.

### Revisió

en aquesta tasca t'és dues columnes que contenen el mateix text: ha de modificar la columna de l'esquerra, i sobretot ha de millorar la fluïdesa del text. si trobas algun error en el text, no dubti en corregir-lo. Si us plau, escriu un text millorat. pots guardar el text i continuar en qualsevol moment que desitgi

Per aquest test no hi ha cap llistat de termes disponible. Si teniu algun dubte consulteu els enllaços que us proveïm a la guia d'estil.

**ATENCIÓ:** és molt important que segueixis les indicacions de la guia d'estil. Et recomanem que imprimeixis el document. El pots fer servir tant durant la prova com més tard durant l'experiment.

-  [Guia d'estil.pdf](#)
-  [Glossary.pdf](#)
-  [File Original.pdf](#)

Millora la frase si cal  
sense canviar el significat



Frase original  
a revisar



Instal·lació de la Passarel·la de mesura (MGT) i Configuració	Instal·lació de la Passarel·la de mesura (MGT) i Configuració
L'aplicació MGT és una passarel·la que metering que accepta dades de mesura des de grids dins del datacenter d'un client i envia aquelles dades al sistema de mesura	L'aplicació MGT és una passarel·la que metering que accepta dades de mesura des de grids dins del datacenter d'un client i envia aquelles dades al sistema de mesura
Verifiqui que cap missatge d'error no es registra al quadre de comandaments de grid.	Verifiqui que cap missatge d'error no es registra al quadre de comandaments de grid.
Si hi ha missatges d'error, sisplau contacti suport tècnic.	Si hi ha missatges d'error, sisplau contacti suport tècnic.

[Guardar](#) [Guardar i Enviar](#)

When user logs in and starts the Review task by clicking the relative button, profile Servlet calls REDetailServlet. This Servlet loads the final text from the PE\_text table (column 'version 2') by looking at the different worker\_id present in the Unit\_id. The Unit\_id at Review phase is like this:

Languageoftheexperiment – PostEdit workerID – Review workerID.

I.e: ct-1-4 ;

Consequently, REDetail Servlet loads text from PE\_text table and saves it in the RE\_text, with the Review worker id.

```
String q1 = "SELECT * from `re_text` where `re_workerid` = '" + re_workerid + "' and `experimentId` = '" + experimentid + "'";
PETranslationManager ptrm1 = new PETranslationManager(q1);

    if ( ptrm1.getLength() == 0 ){
        String q2 = "INSERT INTO `re_text` (`experimentId`,`re_workerid`,`sentenceid`,`pe_workerid`) SELECT `experimentId`,`"+re_workerid+"`,`sentenceid`,`"+
            +pe_workerid+"` from `pe_text` where `experimentId` = '" + experimentid + "' and `pe_workerid` = '" + pe_workerid + "'";

        PETranslationManager ptrm2 = new PETranslationManager(q2);
        ptrm2.closeConnection();
    }
ptrm1.closeConnection();

String q = "SELECT `re_text`.`idre_text`, `re_text`.`experimentId`, `re_text`.`sentenceid`, `pe_text`.`version2`, "
+ "`re_text`.`version1` "
+ "FROM `pe_text` "
+ "left join `re_text` "
+ "on `pe_text`.`experimentId` = `re_text`.`experimentId` "
+ "and `re_text`.`sentenceid` = `pe_text`.`sentenceid` "
+ "where `re_text`.`re_workerid` = '" + re_workerid + "' "
+ "and `pe_text`.`experimentId` = '" + experimentid + "' "
+ "and `pe_text`.`pe_workerid` = '" + pe_workerid + "'";
PETranslationManager ptrm = new PETranslationManager(q);
```

When worker at this phase submits the task, the improved text is saved in the RE\_text table, in column "version 1"

### Table crowd.re\_text

Implied relationships  Related columns  Constraints  Comments  Legend

Column	Type	Size	Nulls	Auto	Default	Children	Parents
idre_text	int	11		√		<i>re_errors.idre_text</i>	
experimentId	varchar	45					
re_workerid	int	11					
pe_workerid	int	11	√		null		
sentenceid	int	11					
version1	text	65535	√		null		
version2	text	65535	√		null		

Table contained 0 rows at Tue Jun 26 20:07 CEST 2012

#### Indexes:

Column(s)	Type	Sort
idre_text	Primary key	Asc

#### Close relationships:

At this point, the Text has been improved by a native speaker of the target language, and it is going to be verified by a group of Verificators at phase Verification of Review.

## **5.9 Verification of Review**

As same way as in the previous verification phase, the text is given to a group of users in charge of checking the work done by one other user, but, during the verification of Review, only monolingual users are involved. Furthermore, users assigned at this task can select among a reduced amount of errors, since all the errors inherent to translation are not present during this phase.

In other words, users at this phase have to take care about the style and focus on coherence of the text produced by the worker at Review phase, finding terminology or grammar errors.

## Revisió

Pren-te el teu temps per tal de contestar a les errors de frases que es plantegen en aquesta pàgina. Una vegada facis clic al botó "Emmagatzema la resposta i continua" ja que no podràs modificar les teves respostes i el sistema emmagatzemarà els resultats a la nostra base de dades. Durant la prova, pots consultar la guia d'estil o qualsevol altra font d'informació. Quant més investiguis millor qualitat del teu treball!

Per aquest test no hi ha cap llistat de termes disponible. Si tens algun dubte consulteu els enllaços que us proveim a la guia d'estil.

**ATENCIÓ:** és molt important que segueixis les indicacions de la guia d'estil. Et recomanem que imprimeixis el document. El pots fer servir tant durant la prova com més tard durant l'experiment.

-  [Guia d'estil.pdf](#)
-  [Glossary.pdf](#)
-  [File Original.pdf](#)

Frase original



Instal·lació i Configuració de la Passarel·la de mesura (MGT)	Instal·lació i Configuració de la Passarel·la de mesura (MGT)
---	---

Opcions de software

Error ortogràfic

Puntuació

Estil

Gramàtica

Terminologia

Deixa el teu comentari aquí

L'aplicació MGT és una passarel·la de mesura que accepta dades de mesura provinents de grids dins del centre de dades d'un client i les envia al sistema de mesura de <productname>.	L'aplicació MGT és una passarel·la de mesura que accepta dades de mesura provinents de grids dins del centre de dades d'un client i les envia al sistema de mesura de <productname>.
--	--

Opcions de software

Error ortogràfic

Puntuació

Estil

Gramàtica

Terminologia

Deixa el teu comentari aquí

Tots els drets reservats.	Tots els drets reservats.
---------------------------	---------------------------

Opcions de software

Error ortogràfic

Puntuació

Estil

Gramàtica

Terminologia

Deixa el teu comentari aquí

Guarda i Envia



When user logs in and accept the verification of Review task, the profile Servlet calls VRDetail Servlet, which loads the improved text from the Review phase saved in RE\_text table (column 'version 1') and the final text from the Post edit phase, which is contained in PE\_text table (column 'version 2'). In this way worker at Verification of Post edit knows the original text from which the worker at Review phase has started to do his modification and improvements. As explained in the design chapter, the texts shown at verification of Review are two: both the original and the improved text. In this way, worker at this phase can check that worker at Review phase did not change the meaning of the translation done at previous phases.

For this reason, an inner query between the PE\_text table and the RE\_text table is done, and both texts are loaded and shown during the Verification of Review phase.

```
String q = "SELECT `re_text`.`idre_text`, `re_text`.`sentenceid`, "
    + "`pe_text`.`version2`, " + "`re_text`.`version1` "
    + "FROM `re_text` " + "inner join `pe_text` "
    + "on `re_text`.`sentenceid` = `pe_text`.`sentenceid` "
    + "and `re_text`.`experimentId` = `pe_text`.`experimentId` "
    + " and `re_text`.`re_workerid` = '" + re_workerid + "'"
    + " and `pe_text`.`pe_workerid`='" + pe_workerid + "'";
ptrm.execute(q);

Translation tr = null;
while (ptrm.moveToNext()) {

    tr = new Translation();
    int id = 0;
    try {
        id = ptrm.getResultSet().getInt("idre_text");
    } catch (SQLException e) {
        id = 0;
        e.printStackTrace();
    }
    tr.setId(id);
    String original = null;
    try {
        original = ptrm.getResultSet().getString("version2");
    } catch (SQLException e) {
        e.printStackTrace();
    }
    tr.setDefaultLanguage(original);
    tr.setVersion1(ptrm.getVersion1());
    list.add(tr);
}
```

As same way as in the previous verification, user at this phase can not save and he can just submit the task. When user submits the task, the errors and comments he has inserted in the interface are saved in the RE\_errors table. When user submits the task, task status is updated from status 'pending' to status 'done' in table REVEUnit, de facto removing the task from user list.

Similarly to the Verification of PostEdit, users can just submit the task and not save it, and errors found at Verification of Review phase are saved in the RE\_errors table.

### Table crowd.re\_errors

Implied relationships  Related columns  Constraints  Comments  Legend

Column	Type	Size	Nulls	Auto	Default	Children	Parents
id_reerrors	int	11		√			
vp_workerid	int	11	√		0		
vp_errortype1	bit	1	√		b'0'		
vp_errortype2	bit	1	√		b'0'		
vp_errortype3	bit	1	√		b'0'		
vp_errortype4	bit	1	√		b'0'		
vp_errortype5	bit	1	√		b'0'		
vp_errortype6	bit	1	√		b'0'		
vp_comment	text	65535	√		null		
idre_text	int	11	√		null		<i>re_text</i>
re_workerid	int	11					

Table contained 0 rows at Tue Jun 26 20:07 CEST 2012

#### Indexes:

Column(s)	Type	Sort
id_reerrors	Primary key	Asc

At this point, the text has been verified by a native user of the target language, and if the text needs some change, it goes back to the worker of Review phase which is in charge to correct them.

## 5.10 Review phase 2

As final step in the translation process, it comes the phase Review 2, where the same worker assigned at Review phase receives comments and feedback from the other users of the crowd, and he can provide the final version of the improved text.

User at this phase sees the sentences which were mistaken or which they can be improved. Worker can see the errors found and read the comments left by other users, and then decide if to change the initial version he provided or just submit the same text he initially provided. The interface used at this phase is similar to the interface at phase Post Edit 2.

When user connects and starts the Review 2 task, R2Detail Servlet is called and, similarly to Post edit phase 2, it loads the text user provided at Review phase, and then it searches for the relatives errors in the RE\_errors table.

```

String q = "select distinct `re_text`.`idre_text`, `pe_text`.`version2`, `re_text`.`version1`, `re_text`.`version2`"
    + "from `pe_text`"
    + "left join `re_text`"
    + "on `re_text`.`experimentId` = `pe_text`.`experimentId`"
    + "and `re_text`.`sentenceid` = `pe_text`.`sentenceid`"
    + "inner join `re_errors`"
    + "on `re_errors`.`idre_text` = `re_text`.`idre_text`"
    + "where `re_text`.`re_workerid`='" + re_workerid + "'"
    + "and `pe_text`.`pe_workerid`='" + pe_workerid + "'";
PETranslationManager ptrm = new PETranslationManager(q);
Translation tr = null;
model.Error error = null;
ErrorManager em = new ErrorManager();

while (ptrm.moveToNext()) {
    tr = new Translation();
    int id = 0;
    try {
        id = ptrm.getResultSet().getInt("idre_text");
    } catch (SQLException e) {
        e.printStackTrace();
    }

    tr.setId(id);
    String original = null;
    try {
        original = ptrm.getResultSet().getString("pe_text.version2");
    } catch (SQLException e) {
        e.printStackTrace();
    }
    tr.setDefaultLanguage(original);
    tr.setVersion1(ptrm.getVersion1());

    try {
        version2 = ptrm.getResultSet().getString("re_text.version2");
    } catch (SQLException e) {
        e.printStackTrace();
    }
    tr.setVersion2(version2);

    translationList.add(tr);

    q = "SELECT `re_errors`.`id_reerrors`, "
        + "`re_errors`.`vp_workerid`, "
        + "`re_errors`.`vp_errortype1`, "
        + "`re_errors`.`vp_errortype2`, "
        + "`re_errors`.`vp_errortype3`, "
        + "`re_errors`.`vp_errortype4`, "
        + "`re_errors`.`vp_errortype5`, "
        + "`re_errors`.`vp_errortype6`, "
        + "`re_errors`.`vp_comment`, "
        + "`re_errors`.`idre_text`, "
        + "`re_errors`.`re_workerid` " + "FROM `re_errors` "
        + "where `re_workerid`='" + re_workerid + "'"
        + "and `idre_text` = '" + tr.getId() + "'";

    em.execute(q);
    ArrayList<Error> tmpErrorList = new ArrayList<Error>();
    while (em.moveToNext()) {
        error = new Error();
        error.setId(em.getIderrors());
        error.setReviewId(em.getVpworker());
        error.setErrorType1(em.getVpErrortype1());
        error.setErrorType2(em.getVpErrortype2());
        error.setErrorType3(em.getVpErrortype3());
        error.setErrorType4(em.getVpErrortype4());
        error.setErrorType5(em.getVpErrortype5());
        error.setErrorType6(em.getVpErrortype6());
        error.setComment(em.getVpComment());
        tmpErrorList.add(error);
    }
}

```

When user submits the task, the text provided is saved in the RE\_text table, in version 2 column. This is also the final output of the crowd translation provided by this prototype.

## 6 Result Analysis

During this project it was decided to run experiments and evaluate different aspects of Translation done in a Crowd sourcing environment. Some of the main things searched in the experiment were the following:

- Potentiality of crowd translation: the quality, speed in translation and coherence of a text translated by multiple persons;
- Provide some data to CA Labs: to support hypothesis that crowd sourcing can be a valuable option for CA Technologies to cope with workload of its Translation and localization department, Heads of CA Technologies requested real data and results.
- Check different combinations of different users: to see which phases influence more the final result, the work of each user at each phases wanted to be analyzed.

To have people registering in the system, experiment have been promoted in different ways, from posters attached around different university campus, word of mouth, Facebook, posts on translators blogs and even by publishing the experiment on the official CA Technologies web pages.

The main motivations behind users participation in the experiment were the following:

- CA Technologies offered the possibility to win an Ipad 2 for all the participants;
- Moral and ethical motivations of being asked to do a favor by a friend;
- Ideal of supporting science by participating an experiment.

Many professional translators also participated with the clear intention of staying updated of the latest trends in translations and understand where the translation world and translation job market are heading towards. For the same

reason, many students from language studies also participated.

Roughly, registered users can be grouped in those main classes:

	<b>Professional</b>	<b>Student</b>	<b>Amateur</b>
<b>Bilingual</b>	51	40	89
<b>Monolingual</b>	10	9	114

Another reason to run some experiments was to show (if exists) any relation among different kind of users allocated to different phases, which may lead to different quality in the final translation. In other words, the scope of the experiment was then to understand if, necessarily, professional translators allocated in the different phases of the designed process, would always perform better than a mix of professional and amateur. Or even if the right combination of amateur would be able to provide high quality comparable to the professional one, with some frequency and reliability.

To create different combinations of different categories of users with different skill sets, there is the need to evaluate users and classify them in different categories. Since the ranking system was not implemented, mainly because the designed ranking system would take some ramp up time in order to correctly classify users. Moreover, users needed to provide a significant amount of work before being classified according to their capability in translation, and this was not possible to ask to users, since they were not earning any money from the work they were providing.

For those and other reasons, users were tested and classified according to results obtained from an initial test meant to evaluate users translation skills.

Test was created by the Translation and Localization department of CA Technologies, and it was proposed to user accordingly to their native language (mainly: Catalan and Spanish) and knowledge of English.

Initial test was created with similar interfaces to the one user finds during the experiment, so they could become more experienced and feel more

comfortable with the system, hence become less prone to errors and slips during the experiment.

Experiment was designed to have users divided into different groups according to the score they obtained; Under the suggestion of the CA translation and localization department, professionals translators were used as benchmark, taking into account that professionals have had training and more experience than the average users, so they can obviously make good translations. Therefore, Any user that would score within a 0,5 standard deviation below the professionals would be put into the group A. Users 0,5 to 1,5 standard deviations below the professionals would be put into group B and people 1,5 to 2,5 would be put into group C. People more than 2.5 standard deviations below the professionals would be discarded for the experiment due to lack of quality. Note that only the mean of the professionals was used as a benchmark, and professionals that scored 0,5 to 1,5 below the reference mean, they would still be placed in the B group (and not A).

## 6.1 Result Analysis

A total amount of 317 users have been registered in the system, up to now. The initial test was proposed as prerequisite to participate in the experiment and, consequently, to run a chance to win the Ipad 2. Eventually, just 150 users completed the initial test, which would take from 30 to 45 minutes.

It can be noticed how meaningful is the clear gap between the amount of registered users, and the users that completed the initial test. In spite of the fact that the following problems occurred during the experiment:

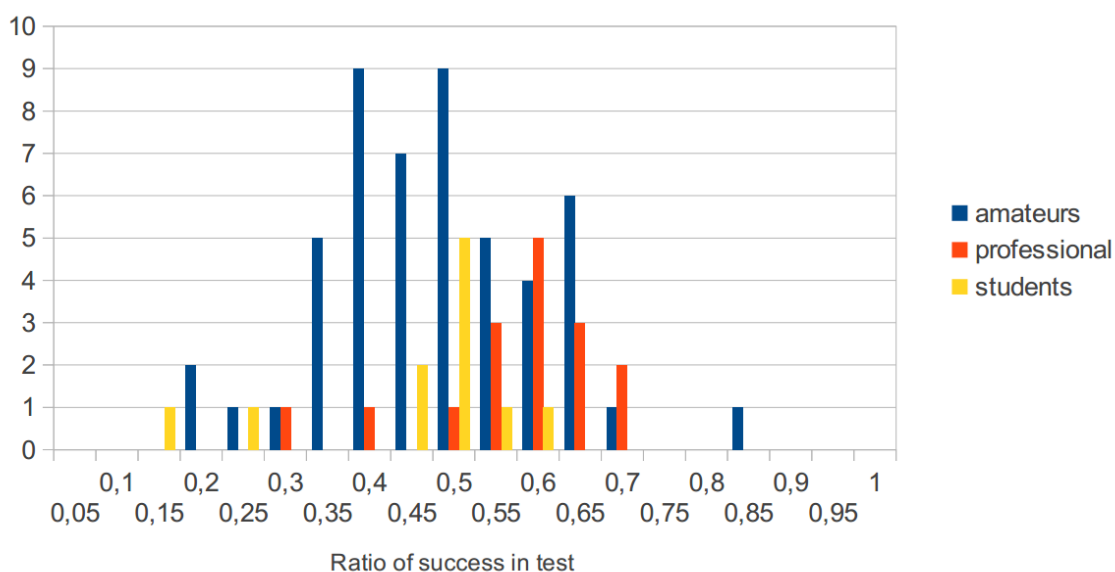
- Too much time passed between the registration period, the initial test period, and the real experiment period. Users may have felt that experiment was proceeding sloppy;
- Initial test was quite hard and really time consuming. Many users started the test just to quit it half way;

This significant gap between registered users and completed test users underlines how motivation to participate in crowd sourcing systems is the key to switch on the engine. Users need to feel that the system is completed, all functional, easy to use and it is fully performant, all at the same time; users need to feel their work is useful, and they have to feel rewarded for their effort.

As said before, the initial test turned out to be more complicated than what it was supposed to be, and few people obtained more than 0,7, consequently, users were grouped differently than what was planned. Group A is composed by users being the top 20% percent, group B being the top 50% to 20% and group C being the top 80% to 50%. The last 20% was discarded as not approaching any of the minimum criteria, or, more likely, those users were not really interested in the experiment and they may have registered just for the Ipad giveaway.



## Histogram



As it can be seen, the Amateur class vaguely follows a normal distribution. From the picture the other two classes apparently show a similar trend, but there are not enough data to claim that. Most of the professionals are concentrated on the "half good part" of the amateur distribution.

Consequently, users have been divided in three main categories: A, B, C, according to the score they obtained in the initial test. Then, users have been selected to be assigned to specific task with Wolfram alpha randomizer.

Catalan								
Post Editors			B			C		
A			B			C		
number of users	27		number of users	27		number of users	26	
number of PE	4		number of PE	4		number of PE	4	
range	6,75		range	6,75		range	6,5	
Row starts at:	7		Row starts at:	7		Row starts at:	7	
5,15			6,62			3,05		
unrounded #	ROW	Worker ID	Unrounded #	ROW	Worker ID	Unrounded #	ROW	Worker ID
5,15185	11	340	6,62177	13	349	3,05405	9	305
11,90185	18	252	13,37177	19	152	9,55405	16	165
18,65185	25	290	20,12177	26	224	16,05405	22	155
25,40185	31	378	26,87177	33	437	22,55405	29	319

(example of Post Editor selected users.)

### 6.1.1 Results

At day 28th of August, unfortunately, the experiment is still in the middle of the designed translation process. Undoubtedly, the experiment took more than the expected time, and experiment is currently at phase Post edit 2 for both Catalan and Spanish users.

In spite of the many emails and reminders that had been sent during the last months, users have not completed tasks due to many reasons such, in example, summer holidays, lack of user's time, task's not trivial difficulties and, probably, lack of motivation, impeding the experiment to proceed smoothly and end within the expected needed. Before starting the experiment, the team expectation was that experiment shall last 6-7 weeks, and the translation should have gone through all the phases of the translation process, so to have some results ready by September.

Anyhow, reader can find attached to this document, some of the translated texts by Catalan users, so the reader may have an opinion by his own about the quality of the translation. Attached texts have been randomly selected, and even if they can not be considered representative of the quality that designed system could provide, those text can be taken as an example of users capability in translation (but not the system performance).

## 6.2 Survey

When users completed their task, the prototype asked them to fill a small Survey. The Survey was created using google docs, and link was shown in user's profile page:

The screenshot shows a user profile page with the following elements:

- Top right: Username: and.gritti@gmail.com
- Navigation buttons: INICI, PERFIL, SORTEIG DE L'IPAD, FAQ, SORTIR
- Logos and text on the left:
  - UNIVERSITAT POLITÈCNICA DE CATALUNYA BARCELONATECH
  - in collaboration with CA Labs
  - funded by FEDER, GOBIERNO DE ESPAÑA, MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD, INVESTIN SPAIN
- Section: Tasques Pendants
- Text: Tens les següent tasques pendants. Fes clic a "Acceptar tasca" per començar a treballar-hi.
- Table:

Tipus de tasca	Estat
Enquesta final	Respon aquestes darreres qüestions per acabar l'experiment
- Button: Acceptar \*
- Footnote: \* Aquest enllaç et portarà a una pàgina externa. Un cop acabis l'enquesta hauràs acabat totes les teves tasques. L'enllaç romandrà en aquesta pàgina, però només cal que facis l'enquesta un sol cop.

The survey was meant to gather user reaction and impression, and investigate user's feelings and approach to a similar system. The magnitude of the survey could have been an analysis of potential users needs and requests, gather of users suggestions, inquiry of users motivations, and it would have been more significant if more users would have been more participative.

Some of the question asked are shown in the following image.

Quant de temps heu trigat a finalitzar les vostres tasques? No cal incloure-hi el temps dedicat a la prova de traducció.	He utilitzat de manera activa els fitxers de suport per finalitzar les meves assignacions.	Teniu cap suggeriment sobre aquesta plataforma?	Estaríeu disposats a treballar en aquest sistema de traducció?	No m'interessa col·laborar com a traductor en aquest sistema per passar l'estona.	Un motiu important per treballar a una plataforma d'aquesta mena seria que milloraria els meus coneixements i habilitats.	Una de les raons importants per les quals he participat és que el meu treball s'utilitzarà en un producte real.	Si rebés una compensació econòmica segons la qualitat de la meva traducció, li dedicaria més temps i esforç per entregar una traducció de més qualitat.	Un factor important per participar en aquesta plataforma ha estat el meu interès en la traducció.	Quantes hores a la setmana treballaríeu en una plataforma com aquesta?
45-60'	Totalment d'acord		Sí	En desacord	Totalment d'acord	Totalment d'acord	Totalment d'acord	Totalment d'acord	3-6 hores a la setmana
30	D'acord		No	D'acord	Indiferent	Totalment d'acord	D'acord	Totalment en desacord	1-2 hores a la setmana
50	D'acord		Sí	En desacord	D'acord	Indiferent	En desacord	D'acord	3-6 hores a la setmana
60	D'acord	Poder guardar les respostes a mig fer i continuar més tard abans d'enviar.	Sí	En desacord	D'acord	Indiferent	Totalment d'acord	Totalment d'acord	1-2 hores a la setmana
30	Totalment d'acord	Indicar terminis d'entrega de les tasques.	Sí	En desacord	Indiferent	Indiferent	D'acord	D'acord	1-2 hores a la setmana
30	Totalment d'acord	Especificar terminis d'entrega de les tasques	Sí	En desacord	Indiferent	Indiferent	D'acord	D'acord	1-2 hores a la setmana
50	D'acord		Sí	En desacord	D'acord	D'acord	Totalment d'acord	D'acord	7-12 hores a la setmana
120	D'acord		Sí	Totalment d'acord	D'acord	D'acord	Totalment d'acord	D'acord	3-6 hores a la setmana
90	Totalment d'acord		Sí	En desacord	Totalment d'acord	Totalment d'acord	Totalment d'acord	D'acord	3-6 hores a la setmana
30,3	D'acord		Sí	En desacord	D'acord	Totalment d'acord	D'acord	Indiferent	7-12 hores a la setmana

(some of the question asked for the survey)

## 7 Comparison

In this chapter I make profit of the knowledge acquired on crowd sourcing to evaluate and compare different crowd sourcing systems. In the first part I take two papers as reference to rattle off different aspects and criteria used to compare the designed system with others similar competitors such it may be Amazon Mturk, Duolinguo and Icanlocalize. In the first paper[5], which has been already mentioned in chapter 3, authors have extensively studied many different definitions of crowd sourcing, looking for common aspects and trying to create an unique and correct definition of what is crowd sourcing. This paper lists many aspects and characteristics which I used as metrics to compare the designed system with some of the systems described in the previous chapters.

In [26] authors have studied different crowd sourcing processes used to obtain and aggregate contributions from the crowd. Greiger at all have identified four dimensions that describe how crowd sourcing processes differ. Those four dimensions are shown in the following img.

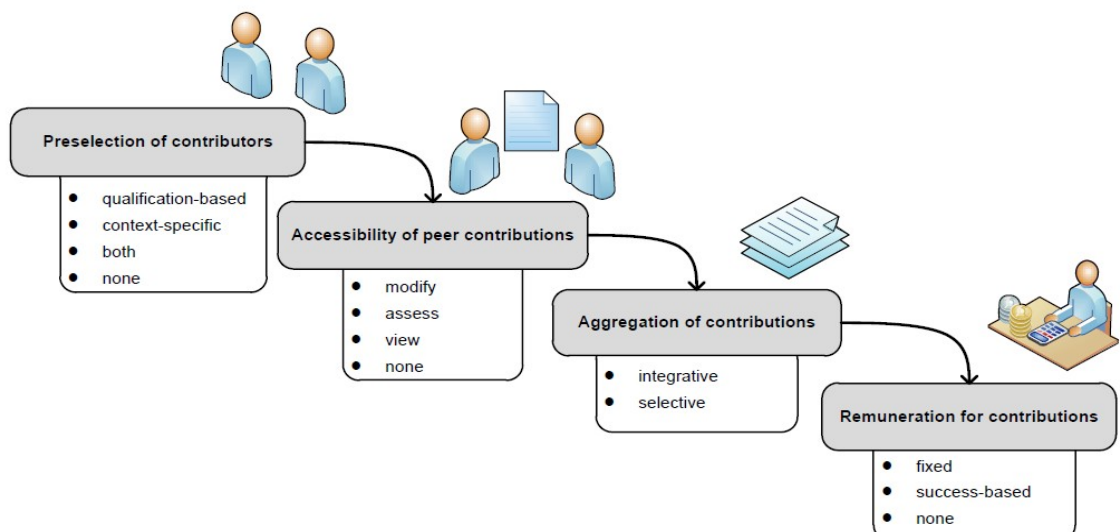


Figure 2. Characteristics of crowdsourcing processes

In particular, Authors addressed the fact that companies striving to achieve a specific goal, should be able to evaluate the mechanisms that impact crowd sourcing processes, when, nowadays, there are not clear gauge for judgment.

For this reason, I tried to address this issue by formalizing the procurement process in using those systems by suggesting a simple road map to follow, in order to decide which crowd sourcing system is the most suitable for the problem to be solved. To do this, I have followed the example of the procurement of another, similar, technology is described, namely, the acquisition of cloud computing [27].

## 7.1 Discussion

### Who forms the crowd.

The designed system does not pose any preselection of contributors. Every kind of user can be registered in the system and participate in a task. Even if there is the possibility that ranking module may let some users with low ranking to starve by not assigning any task to them, ideally, all the users should be able to participate, contribute, translate and earn some money, especially if the amount of text to be translated in the system is high, and the targeted languages are many. Intuitively, most popular languages will need more users, since the amount of text would be bigger, but even users speaking minor languages can be needed. The system does not do any preselection of user, neither based on qualification, such as 'Icanlocalize', or context-based, as it is done in 'Amazon mechanical turk'. Since bilingual and monolingual are used during the translation process, all users can potentially be involved. Logically, users are assigned to tasks they can perform and about languages they master.

Therefore, the distribution of participants would not be totally random as it shall be in Amazon mechanical turk, because the designed system is specialized on translation (as it is now), and users can perform just tasks about translation. On the other hand, the system does not need strictly specific distribution of user, as it may be in other crowd sourcing system such, in example, 'Innocentive'\* or 'Starmind\*\*', where the crowd present in those systems is made up to more than 66% of PhD student, postdoctoral, researcher and in general, more educated users, due to the high difficulty of the tasks.

In other words, the crowd needed in the designed systems is not limited to professionals such as in Icanlocalize.com, nor the designed system is appealing just for a group of people that are interested in learning new languages, such in 'Duolingo'.

\*[[www.innocentive.com](http://www.innocentive.com)]

\*\*[[www.starmind.com](http://www.starmind.com)]

Undoubtedly, the system would call the attention of Professional translators and persons with interest in translations, since those users, supposedly, are more prone to participate in such a system, especially because they would increase their ranking faster than normal users, consequently earning more money. Anyhow, the system is designed in a manner where there are few barrier to enter (be able to speak a language, willed to provide good quality work), and where each user has a chance to do some earnings.

*What it has to do, which tasks are present in the system*

In principle, any non-trivial problem can benefit from crowd sourcing. Hence, Translation is adequate as crowd sourcing task, and the many studies and researches done and presented in the previous chapters, support this idea. At the same time, crowd sourcing tasks ought to be divisible into lower level tasks to obtain better results [28]. The objective of the designed architecture is not the mere translation, but it is a methodological assurance of high quality in the final translation. Therefore, the translation process has been divided in smaller tasks, each one with a clear objective. This methodology is novel and unique, it comes from the traditional process to translate text, and it has been adapted in a distributed system; there is no similar system that offers this methodology in an automated and fully managed manner, ergo comparison is not possible. After all, the same tasks can be also done on 'Amazon Mechanical Turk', but it would need a manual organization of all the steps, which is not suitable for a large and distributed system. At the same time, 'Duolingo' approaches user's needs to learn the new language first, before he can really be productive and deliver good quality translation. Even if the idea to motivate user by claiming they will learn a new language is alluring, the quality of the final translation won't be high with a certain frequency, especially because users are not mastering the new language. For this reason, 'Duolingo' is focusing on translating the web, and is not thought for an industrial utilization where certain quality standards have to be fulfilled. 'Icanlocalize' has probably an internal



organization which is similar to the translation process created in this project, but as it is said before, 'Icanlocalize' is more a community of translators rather than a crowd sourcing system.

#### Accessibility of peer contributions

In the designed system, users interact with other users by detecting errors and providing feedback and access to the material created by others accordingly to the Action-Verification pattern described in the previous chapter. Crowd sourcing has been used to create information or to vote and select the best among a group of information, but those two different tasks have never been related each other in a methodological path. A similar system has been proposed in [29] where a “control group” is used to detect cheater, more than to improve the quality of the product provided by another user.

#### Aggregation of contribution

The designed system use the Map Reduce paradigm to aggregate the contribution of each user in order to create the final translation. Therefore all contributions are used for the final outcome; it may be that for some contributions, more than one execution is needed, but all the final contributions are needed, since all of them are different parts of an entire document. In contrast to other systems, there is not any selective contribution where the best translation is chosen to become the final one, as it is in 'Duolinguo' for example.

#### How crowd is rewarded

The proposed system is designed to be able to create the possibility for users to earn some real money. Since the produced material, as it may be a manual for a software system, would be used and sold with the product itself, micro-payment was not even taken into consideration, but rather the cost of the current translation market was taken as a reference. Moreover, the ranking

system is thought to reward the experience and time spent on the system, as same as the quality of the work produced. In comparison to micro-payment offered in 'Amazon mechanical turk', the rewarding designed method seems more legitimate and fair. Web sites as 'Icanlocalize' offer translation and localization at really low prices, which are around 10 cents per word, rising up to 15 cents per word for Translation and a Review. Consequently, those prices are the amount of money the designed system should compete with. While this amount of money may seem the minimum price to pay for professional translators, this amount of money may be satisfying for non professional translators, native speakers and, in general, amateurs.

#### *Who is the initiator, the stakeholder*

In the designed system, anyone can request the translation: SMB, big companies or even single users. The initiator is common to many different systems, since the need for crowd sourcing way for problem solving can be pretty general.

Some systems such as 'Duolinguo' are privately used by their creators for known (translating the web) and, maybe, others unknown objectives.

#### *What the stakeholder gets in return for the work of the crowd.*

The designed system is totally focused on translation. Even though the Action-Verification pattern can be generalized for other tasks, the system has been created for translation processes. Consequently, up to now, the system can only offer high quality Translations services. In comparison, 'Amazon mechanical turk' is not implemented specifically for any particular task, and it is actually used for a variety of different daily grind. On the other hand, the obtainable outcome that 'Amazon mechanical turk' can provide, is usually not satisfactory or just sketchy, and often useless due a scarce quality, resulting in a waste of time and resources.

*What type of process it is.*

The designed process automatically changes and adapts the translation process, accordingly to users selected for the translation. Any other mentioned system does not have this capability of user management, making the designed system unique under this criteria. Users contribute on line via common web technology, while in other systems such 'Icanlocalize', translators shall work offline by downloading the texts and uploading it once completed.

*The type of call used.*

While in many other crowd sourcing systems, the tasks are offered as an open call, the designed system works the other way around. When users set their status as 'available', is the system which offers to users the task, with the relative deadlines to accept the task and deliver the work. This methodology is in contrast with 'Amazon mechanical turk', where users apply to tasks, but it is similar to 'Duolinguo' and other systems, where it is the system that proposes tasks to users, and they can accept to do it or not.

	<b>Designed System</b>	<b>Amazon Mechanical Turk</b>	<b>Duolinguo</b>	<b>Icanlocalize</b>
<b>Crowd composition</b>	+ Specific and general kind of users	++ totally general kind of users	- users that want to learn new language	-- professional translators
<b>Kind of tasks</b>	- translating and reviewing	++ Any kind of task	- learning and then translation	- just translating
<b>Reward for crowd</b>	++ money according to quality	- few money	+ possibility to learn new language	+ money for a professional service
<b>Who is the requester</b>	- whoever need a translation	+ any one	-- owner of the system	- whoever need a translation
<b>What the requester obtain</b>	++ quick high quality translation	- simple tasks get done	- translation of web	+ professional translation
<b>Type of process</b>	++ automated, adaptive	-- simple and linear	+ long	+ offline

## 7.2 Crowd Sourcing Adoption

In this chapter, I provide some guidelines for crowd sourcing procurement. Due to totally lack of literature on this field, I use literature from other more traditional technology procurement, as it may be ERP, and some recent literature on Cloud computing adoption, which it has some common elements with crowd sourcing, but where I also highlight some intrinsic differences between those two recent technologies. Moreover, traditional outsourcing literature is also mentioned, since, crowd sourcing, it is a innovative methodology which blend Cloud computing and traditional outsourcing.

### Task Suitability Analysis:

Crowd sourcing is not cloud computing. While cloud computing is said it will bring a shift in IT, crowd sourcing is not compulsory. Therefore, companies may decide to adopt or not this technology, while some tasks usually done internally to the company are more suitable to be crowd sourced. So, the first thing to approach crowd sourcing is to gather all the time consuming tasks which do not involve any sensitive information or high skilled worker. Basically, crowd sourcing procurement works the other way around to other technology procurement: the first step is not an analysis of technology suitability, but more an analysis of task suitability.

Task Suitability Analysis comprises a simple checklist of questions to provide a rapid assessment if a specific task can, potentially, be crowd sourced. The checklist, shown in the following table, analyzes different characteristics of the task, and quickly provides an indication if the task is suitable to be crowdsourced.

FIELD	QUESTION
Employee's role	is the training needed for this task extensive?
	is the employee's experience a key factor for the success of the task?
Task's characteristic	Is the task time consuming for internal employee?
	Can the employee currently working on this task, be assigned to a more productive task?
Data	Are the data involved in the process highly sensitive ?
	Are the data digitized, easy to distribute and manage?
Work management	Are deadlines strict for the task?

The outcome of the analysis is a recommendation of whether or not to proceed with further analysis.

#### Crowd system selection:

A study to find out which crowd sourcing system suits the best, according to the features shown in the chapter above and the characteristics of the task. Then the cost calculation of its use has to be done. Once the system has been selected, initiator have to decide how to propose the task to crowd: open auction for the cheapest solution, participation opened to everybody, collaborative work, challenge to win.

#### Responsibility assignation:

Since Crowd system use may help different departments, it must be identified the potential viability of its use, by mapping responsibilities in order not to leave new responsibility uncovered and not accountable.

### 7.2.1 Cost Calculation

While understanding the operational costs of public clouds may result complicated because of its hidden cost, crowd sourcing presents an easier calculation, but it still has some not obvious problems which may result in revenue loss.

The use of crowd sourcing systems increases difficulty to management when it comes to strict deadlines, due to an increased uncertainty of the final quality results, the freedom of crowd worker to resign from the task at any time, and, consequently, the amount of money needed to obtain a result may also vary. In example, if a task is paid just if the result is useful for the requester, this allows to fix a maximum budget, but that deadline can easily change due to a constant repetition of the task until quality standards are reached. On the other hand, using a system as the one designed for this project, may assure the quality of the final product satisfies the initiator, but money needed for the entire process may increase due to the high ranking of the users involved.

Another hidden cost of crowd sourcing, as it is in every outsourcing, is the fact that competencies are taken out from inside the company; anyhow, crowd sourcing is often used for trivial tasks (especially at first stages of its adoption) and this shall become a problem just for a massive use of crowd sourcing to perform important tasks.

In other words, with a typical crowd sourcing system, it is possible to decide a fixed budget to allocate for a task to be crowd outsourced. On the other hand, there is the possibility to completely waste the money in case that: the task, the crowd sourcing system and the control parameter are not chosen thoughtfully. The uncertainty relates to: i) the task may be too complicated to be addressed by crowd, resulting in a poor quality result; ii) increased complexity in defining deadlines and delivery dates; iii) traditional problem of outsourcing, where crowd sourcing provider's pricing scheme may change anytime. All those problems may be mitigated with a system as the one that have been designed during this project, especially when it comes to quality assurance.

On the other hand, crowd sourcing offers the typical advantages of outsourcing (not hiring employees and avoiding the relative cost of interns, shift from capital, employees, to operational costs ... ) and increases them to a higher level in terms of flexibility, its adaptive and elastic features make it more adaptable to changes in workload.

Crowd sourcing may free internal employees that can be used for all the tasks which are not suitable to be crowd sourced, and it also would outsource some operational costs such, in example, energy to run an office.

### 7.2.2 Organizational Change

When tasks are outsourced from companies, employees usually feel threaten. Even during this project, during the meetings held with the translator employees of CA Technologies, this apprehension was palpable in the room. The use of crowd sourcing has to be gradual and incremental, meanwhile, company has to clarify what is the objective of the outsourcing, and how the crowd system will be used, so employees can be correctly informed on what are the changes it will bring. Project management would also change, since head of department will have to deal with work done from a crowd of users with, basically, no accountability. Crowd sourcing differs to traditional outsourcing in many things, and it ought to be seen by employees as a tool to extend department capacity and productivity, rather than a totally different outsourced system where to pass over extra work.

The objectives of Crowd sourcing should be the same as in traditional outsourcing: a desire to focus on the core business of company by outsourcing secondary tasks, cutting cost, making a better use of internal resources, and solving operational weakness in a specific department.



## 8 Conclusion

This chapter summarizes what was done and learned during these 6 months project. In addition, some thoughts and ideas I had during the project development are inserted as a personal opinion and suggestion.

The project started as a collaboration with CA Labs in February 2012. It begun by addressing the current problems in the Translation and Localization department of CA Technologies, which are the limitations and impediments that are tampering with CA Technologies business. During this period I was able to collaborate productively within a team and understand the actual situation and environment for the development of the project.

The main topic of this project is Crowd sourcing, which it has been analyzed as a viable option to help CA Technologies Translation and Localization department to cope with burst in the workload. A complete architecture of a crowd sourcing platform that satisfies CA Technologies prerequisites has been proposed after two months of work, and a novel crowd sourcing framework for quality assurance has been also invented.

In addition, a prototype of the designed architecture has been implemented, followed by the definition and execution of an experiment with a limited crowd. While the experiment is still ongoing, the results up to those days have been collected and some text obtained from the systems resulted to be promising.

The result set from the experiment partially shows how quality in translation can be obtained from a crowd of unknown users. The reduced amount of data cannot allow the team to claim that the created methodology of Action-Verification assures quality in crowd sourcing systems with high frequency, but further experiment will be done, and the team is confident that the features invented during this project make the designed systems a valuable corner stone for crowd sourcing system, especially for system focused on some particular tasks as it may be translation.

In my personal opinion, this project has also exposed a sensible nerve of the Translation and Localization world: using native people as translators solve the problem of the various subcultures that coexist inside a single nation made by different sociocultural aspects.

By way of explanation, almost all multinational companies actively present in the Catalonia market, offer their product translated and localized in Spanish language; mostly because translation to Catalan is seen as too onerous and with a limited economic turnaround. Concretely, during the experiment, Catalan users were translating from English directly to Catalan, showing how the translation to less popular languages, as it may be in example Catalan, it can be addressed within reduced cost, if done using a system as the one described in this document.

With the experience gained during this project, some different crowd sourcing systems have been compared, and an intuitive road map in crowd sourcing procurement has been also proposed.

## 8.1 Advices and thoughts

As a personal opinion, Crowd sourcing can be approached from two different prospective: Become a crowd sourcing utilizator or a crowd sourcing provider.

As far as I am concerned, the experience gained during this months makes me believe that CA Technologies could consider crowd sourcing as a valuable option to support its own Translation and Localization team during the peak of workloads. In addition, crowd sourcing can be a useful tool for translating to minor languages to which there are scarcity of translators and the company is not translating nowadays.

As it is shown in the paper [30] the decision between keep a product globalized or make it more local brings different results when it come to sell products in specific markets. For this reason, having a crowd sourcing platform for translation would allow CA Technologies to pierce peculiar markets in a smoother manner, and, in addition, it would also allow to increase revenue in some specific markers, as it may be East Europe where products are nowadays sold in a foreign language of the national one.

From a totally different prospective, CA Technologies can even consider the creation of its own crowd sourcing platform as a new business opportunity to integrate to its already vast range of offered products and services. Due to its flexibility and scale, Translation performed within crowd sourcing system similar to the one presented in this document could be easily offered as a Service, opening the doors to small and medium companies to a professional translation, when they did not have the resources to afford a complete translation and localization of their products. On the other hand, even companies of the same size as CA Technologies may be interested in using such a platform any time the workload increases suddenly, de facto obtaining a share of the market kept by the traditional translation outsourcing companies.

Having access to a professional translation service, offered with the typical "as

Service" characteristic of flexibility and reduced cost of procurement, would support the integration of economy interchange among different countries, due to an increased ease of product localization. Investing in such a project would also be aligned to the recent extensive investment CA Technologies have done on cloud computing.

In case a system as the one described in this document would be build and it would reach the point of having a crowd sourcing platform populated with a significant amount of users, It has been proved as anyone should never look at crowd as cheap labor workforce if they want to obtain quality.

Jeff Howe said:

"crowd sourcing works best when an individual or company gives the crowd something it wants. Another way of thinking about this is successful crowd sourcing involves satisfying the uppermost tier on Maslow's hierarchy of needs. People are drawn to participate because some psychological, social, or emotional need is being met. And when the need isn't met, they don't participate. What this means for companies is they must reverse the thinking that normally goes into employee relations. If iStockphoto had approached community building by trying to create a low-paid workforce of amateur photographers, it would have failed. Instead, founder Bruce Livingstone set out to create one on the Web where enthusiasts could share and critique one another's work and, oh yeah, maybe make a few bucks on the side"

Possibly, if a specific company could understand what brings people to participate in a crowd sourced platform for translations, and then match their personal motivations, this would get the ball rolling.

On the other hand, not everyone gets involved in a project for the same reasons and it is important to ensure there is a variety of different rationales available. Furthermore, the kind of offered tasks should never be too long, since easier tasks will increase the likelihood of contributing or participating if what is required is straightforward and can be fitted into a few minutes.

Crowd sourcing's beauty resides also in the fact that people can participate

during their spare time.

The sense of reward is extremely important. As it is shown in [31] users are more productive and do better quality tasks whenever they feel their work has been evaluated, judged or at least seen by others.

One of my role during this project was the development of the web interfaces where users could perform the different tasks. During the designs and the code of the prototype, I have had the possibility to improve my knowledge on GUI development and human computer interaction. Under the expert eye of the head of the Translation and localization department of CA Technologies in Barcelona, Patricia Palladini, I tried to recreate in users the same mental schema proposed by Trados SDL tool, so to attract professional translators, which have been using Trados SDL tool since years. At the same time, I also made interfaces to be more minimalistic, thus more suitable for crowd sourcing tasks and capacity of the average user. Different small tricks have been put in the interface to ease user work as much as possible. Anyhow, some feedback have been received from users, and a summative evaluation of the interfaces clearly exposed that users needs more support, especially when it comes to technical translations. Users demanded easy access to past translations, in order to use them as an example, and a wiki would also be an useful tool to solve doubts.

## 8.2 Ranking module suggestion

I want to dedicate an entire chapter to provide suggestion over the Ranking module because I consider it as the cerebellum of the entire system.

Ranking module is the one in charge of managing the ranking of workers, and its role is vital for the performance of the entire system. The Ranking system should be complex enough to correctly profile users, and specifically it should not just be able to define the skill set own by a user, but it ought to determine the "Professionalism" of that worker. Professionalism of worker is more important than a potential talent in translation, which it reaches its full capacity few times.

Many more criteria can be added to the actual design of the Ranking module to more precisely tune it, and many of those criteria may be proved right with a real system where a more extensive testing can be performed. Here I try to suggest a list of possible criteria that should be taken into consideration to improve ranking module performance:

1 -- Acceptance task rate: This criteria is meant to analyze the user's will to work, and also it would indicate how much hungry for tasks is the user. Moreover, if the acceptance ratio is related to the date of registration of a user, and this criteria would increase worker's ranking as an attribute which may be called "user experience", this would discourage users to have multiple accounts.

2 -- Cancellation rate: Whenever a user resign from a task, this imply a significant loss of time and computational waste to rearrange the translation process. Even during the experiment, lots of problems were coming from users lately informing us that they could not do the task. This criteria wants to delineate if user has a disposition to accept tasks and do not complete them;

3 -- Punctuality in deliver work: The system needs tasks to be completed in time, to better orchestrate the distribution of work and the translation process.

So, punctuality in deliver work should be taken into account, or even system should give some extra points to users that deliver the work beforehand;

4 -- Language skills: Professional translators should be able to increase their ranking faster, if they want to. Calibrated and increasingly difficult tests, exams and certifications should allow users to faster increase their ranking. As already mentioned in this thesis, if something similar to a "CA Crowd Translator Certification" would become a standard in the translation market, this would increase CA crowd sourcing platform's popularity;

5 -- Quality of work in time: If just skills of user, reputation, or historic of work, are taken into consideration to decide not to check user's work any more, this is not enough to guarantee user will always provide a good work. This is why system should check the quality of each work the user perform quite often, and this is why some automatic quality checking measure should be implemented in the system. Quality measure may be more complex and useful than BLUE or METEORE score;

A more complete set of criteria leads to a better profiling of the user, and this would allow the system to more precisely trust user when assigning him tasks, and generally perform better and deliver better results when it comes to save money by reducing the translation process. Even speed in translation would benefit from a well tuned ranking system, by avoiding frequent pitfalls and rearrangement in the translation process. The ranking method proposed in the designed architecture is, as it is now, mostly focusing on the quality of the work provided, and it is not exactly profiling user behavior.

### **8.3 Knowledge acquired and Reflections**

During the project I have been able to study the potentiality of crowd sourcing, and to understand how crowd sourcing could offer, and how it can be exploited in a more productive way than how it is used right now.

I feel myself a lucky student, having the possibility to collaborate with an important company such as CA Technologies, and I have also been lucky enough to work with stimulating colleagues on challenging topics.

I am also happy I could improve my programming skills, which were bit rusty, and I feel satisfied that I have been able to collaborate profitably with other students to create something valuable.

As a final and personal opinion, when I look back to the begin of this project, I certainly see myself grown in the subject. Even though I have some afterthought on some aspects of the project, I am glad I had the possibility to participate to such as an innovative and visionary idea as it is the use of crowd for translation, and I am partially satisfied of the results. When I say partially I mean because I believe crowd systems could deliver impressive results, and unfortunately, there is a lack of those in this document.

I feel crowd sourcing translation is still far to be completely exploited, and its capability are not fully understood yet.



## **8.4 Future works**

As an individual and as a student, I am still very motivated to keep working on crowd sourcing and realize other ways to exploit it, in order to improve traditional processes and manners to get things done. I will surely support the project and the experiment as long as I will be needed, and I am very curious to see the consequences this project may have.

What I would really like to see is the implementation of the designed architecture with all its functionality working at their full potential. I personally think the designed system has important features which are missing in competitors. Thereupon I would be glad to collaborate with CA Labs in the development of such a system.

In any case, I consider crowd sourcing as part of my personal background and CV, and I am well inclined to deep my knowledge in this area.

I feel that crowd sourcing is still at the expectation rampage of Gartner hype cycle, and, in my opinion, the main shift it may bring in outsourcing methodology, it is still far, and the best of crowd sourcing is yet to come!

## 9 Glossary

ITC = Intertask controller

MC = Map controller

DOCUMENT = object to be translated

PE = Post edit TASK

VEPE = Verification of PE TASK

RE = Reviewing TASK

VERE = Verification of RE TASK

PE worker = worker assigned to PE

VEPE worker = worker assigned to VEPE

RE worker = worker assigned to RE

VERE worker = worker assigned to VERE

## 10 Bibliography

[1] Kano, Nadine. *Developing International Software for Windows 95 and Windows NT. A Handbook for International Software Design*. Microsoft Press: Redmond, 1995. (Chapter 2: Designing a Global Program)

[2] Keniston, Kenneth. *Software Localization: Notes on Technology and Culture*. 1997. [Online Available: <http://stuff.mit.edu/people/kken/papers1/Software%20localization.htm>]

[3] Yuen, Man-Ching; King, Irwin and Leung, Kwong-Sag. *A survey of crowd sourcing systems. In Proceedings of the IEEE Third International Conference on Social Computing (SocialCom)*, IEEE, October 2011. Pages 766–773.

[4] M. Vukovic. Crowdsourcing for enterprises. *Services, IEEE Congress on*, 0:686–692, July 2009.

[5] Estellés-Arolas, Enrique and González-Ladrón-De-Guevara, Fernando. *Towards an integrated crowd sourcing definition. J. Inf. Sci.* 38, 2. April 2012, 189-200.

[6] Paolacci, Gabriele; Chandler, Jesse and Ipeirotis, Panagiotis G. *Running experiments on amazon mechanical turk. Social Science Research Network Working Paper Series*, July 2010.

[7] Smucker, Mark D. and Jethani, Chandra P. *Human performance and retrieval precision revisited*. In *Proceeding of the 33rd international ACM SIGIR conference on Research and development in information retrieval*, SIGIR 2010, New York, pages 595–602.

[8] Yan, Tingxin; Kumar, Vikas and Ganesan, Deepak. *CrowdSearch: exploiting crowds for accurate real-time image search on mobile phones*. In *Proceedings of the 8th international conference on Mobile systems, applications, and services (MobiSys 2010)*. ACM, New York, 77-90.

[9] Wolfson, Stephen M. and Lease, Mathew. *Look before you leap: Legal pitfalls of crowd sourcing* Proceedings of the American Society for Information Science and Technology, 2011. Volume 48, Issue 1, pages 1–10.

[10] Thies, W., Ratan, A., and Davis, J. *Paid Crowdsourcing as a Vehicle for Global Development*. CHI Workshop on Crowdsourcing and Human Computation, (2011)

[11] Chandler, Dana and Kapelner, Adam. *Breaking Monotony with Meaning: Motivation in crowd sourcing Markets* University of Chicago Becker Center on Chicago rice Theory 5807 South Woodlawn, STE 370 Chicago, IL 60637 The Wharton School of the University of Pennsylvania.

[12] Downs, J. S.; Holbrook, M. B.; Sheng, S. and Cranor, L. F. *Are your participants gaming the system?: screening mechanical turk workers*. In *Proceedings of the 28th international conference on Human factors in Computing systems*, CHI '10, pages 2399–2402, New York, 2010.

[13] Callison-Birch, Chris. [Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk](#). *EMNLP-2009: proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, Singapore, 6-7 August 2009; pp.286-295. [PDF, 289KB]

[14] Papineni; Roukos, Salim; Ward, Todd and Zhu, Wei-Jing. *BLEU: a Method for Automatic Evaluation of Machine Translation* Kishore IBM T. J. Watson Research Center Yorktown Heights, NY 10598, USA. Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL), Philadelphia, July 2002.

[15] Zaidan, Omar F. and Callison-Burch, Chris. *Crowd sourcing Translation: professional quality from Non-professionals*. In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies (June 2011), pp. 1220-1229.

[16] Denkowski, Michael and Lavie, Alon. *Exploring normalization techniques for human judgments of machine translation adequacy collected using Amazon Mechanical Turk*. In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT 2010). Association for Computational Linguistics, Stroudsburg.

[17] Bentivogli, Luisa; Federico, Marcello; Moretti, Giovanni and Paul, Michael. *Getting Expert Quality from the Crowd for Machine Translation Evaluation*, Proc. of the MT Summit XIII, Xiamen, China, 2011, pp. 521-528,

[18] Brown, Peter F.; Della Pietra, Vincent J.; Della Pietra, Stephen A. and Merce, Robert L. *The mathematics of statistical machine translation: parameter estimation*. *Comput. Linguist.* 19, 2 (June 1993), 263-311.

[19] Gao, Qin and Vogel Stephan. *Consensus versus expertise: a case study of word alignment with Mechanical Turk*. In Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk (CSLDAMT 2010). Association for Computational Linguistics, Stroudsburg, PA, USA, 30-34.

[20] Negri, Matteo; Bentivogli, Luisa; Mehdad, Yashar; Giampiccolo, Danilo and Marchetti, Alessandro. *Divide and Conquer: crowd sourcing the Creation of Cross-Lingual Textual Entailment Corpora*. Proceedings of EMNLP-2011, Edinburgh, UK, 27–31 July 2011.

[21] Bernstein, Michael S.; Little, Greg; Miller, Robert C.; Hartmann, Bjorn; Ackerman, Mark S.; Karger, David R.; Crowell, David and Panovich, Katrina. *Soylent: a word processor with a crowd inside*. In Proceedings of the 23rd annual ACM symposium on User interface software and technology (UIST 2010). ACM, New York, NY, USA, 313-322.

[22] Ross, Joel; Irani, Lilly; Silberman, M. Six; Zaldivar, Andrew and Tomlinson, Bill. *Who are the crowdworkers?: shifting demographics in mechanical turk*. In *Proceedings of the 28th of the international conference extended abstracts on Human factors in computing systems (CHI EA 2010)*. ACM, New York, 2863-2872.

- [23] Ipeirotis, Panos. *Demographics of Amazon Mechanical Turk*. Department of Information, Operation, and Management Sciences Leonard N. Stern School of Business, New York University.
- [24] Dean, Jeffrey and Ghemawat, Sanjay. *MapReduce: Simplified data processing on large clusters*. In *OSDI 2004*, pages 137–150.
- [25] V. Muntès, P. Palladini, J. Manzoor, A. Gritti, J. Larriba, F. Mijnhardt, *Crowdsourcing for Industrial Problems*, CitiSen 2012. Montpellier, France.
- [26] Geiger, D. Seedorf, S. Schulze, T. Nickerson, R. *Managing the Crowds: Towards a Taxonomy of Crowdsourcing Processes*, In *Processing of the seventeenth Americas Conference on Information Systems*, 2011.
- [27] A. Khajeh-Hosseini, D. Greenwood, J.W. Smith and I. Sommerville, *The Cloud Adoption Toolkit: Supporting Cloud Adoption Decisions in the Enterprise*. Software: Practice and Experience, 2011
- [28] Vukovic, M.; Mariana, L. and Laredo J. *PeopleCloud for the Globally Integrated Enterprise*. In: D. Asit et al. (eds) *Service-Oriented Computing*. (Springer-Verlag, Berlin/Heidelberg, 2009)
- [29] Hirth Matthias, Tobias Hoßfeld, and Phuoc Tran-Gia. 2011. Cost-Optimal Validation Mechanisms and Cheat-Detection for Crowdsourcing Platforms. In *Proceedings of the 2011 Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS '11)*. IEEE Computer Society, Washington, DC, USA, 316-321.

[30] Sangeeta Ramarapu; John E Timmerman; Narender Ramarapu, Choosing between globalization and localization as a strategic thrust for your international marketing effort *Journal of Marketing Theory and Practice*; Spring 1999;

[31] Bernardo A. Huberman, Daniel M. Romero, and Fang Wu. 2009. Crowdsourcing, attention and productivity. *J. Inf. Sci.* 35, 6 (December 2009), 758-765.