

DELIVERABLE

Project Acronym: PLUTO
Grant Agreement number: 250416
Project Title: Patent Language Translations Online

Deliverable 7.3 Final Survey Results

Authors:

John Tinsley (DCU)
Joeri van de Walle (Crosslang)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	
C	Confidential, only for members of the consortium and the Commission Services	x

REVISION HISTORY AND STATEMENT OF ORIGINALITY

Revision History

Revision	Date	Author	Organization	Description
1	20/03/2013	J.Tinsley	DCU	Document created
2	28/03/2013	J.van de Walle	DCU	Translation of results
3	01/04/2013	J.Tinsley	DCU	Final document
4	02/04/2013	J.Tinsley	DCU	Review, copy-editing, and formatting

Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Table of Contents

1	INTRODUCTION	ERROR! BOOKMARK NOT DEFINED.
2	MACHINE TRANSLATION DEVELOPMENTS IN YEAR 3	ERROR! BOOKMARK NOT DEFINED.
2.1	TRANSLATION QUALITY	ERROR! BOOKMARK NOT DEFINED.
2.1.1	<i>Pre-ordering Input</i>	Error! Bookmark not defined.
2.1.2	<i>Japanese Script Normalisation</i>	Error! Bookmark not defined.
2.1.3	<i>Japanese Transliteration</i>	Error! Bookmark not defined.
2.1.4	<i>Chinese Vocabulary</i>	Error! Bookmark not defined.
2.2	TRANSLATION SERVICES.....	ERROR! BOOKMARK NOT DEFINED.
2.2.1	<i>Improved Keyword Extraction</i>	Error! Bookmark not defined.
2.2.2	<i>Alternate Translations</i>	Error! Bookmark not defined.
2.2.3	<i>Unknown words and suggested translations</i>	Error! Bookmark not defined.
2.3	TM/MT INTEGRATION ENHANCEMENTS	ERROR! BOOKMARK NOT DEFINED.
2.3.1	<i>Word alignment information</i>	Error! Bookmark not defined.
2.3.2	<i>Translation Scores</i>	Error! Bookmark not defined.
2.3.3	<i>Probabilistic Bilingual Dictionaries</i>	Error! Bookmark not defined.
3	3 YEARS OF MACHINE TRANSLATION IN PLUTO	ERROR! BOOKMARK NOT DEFINED.
3.1	INTEGRATION AT THE EPO	ERROR! BOOKMARK NOT DEFINED.
3.2	LANGUAGE COVERAGE	ERROR! BOOKMARK NOT DEFINED.
3.3	IPTRANSLATOR	ERROR! BOOKMARK NOT DEFINED.
3.4	TRANSLATION MEMORY INTEGRATION	ERROR! BOOKMARK NOT DEFINED.
3.5	TRANSLATION API	ERROR! BOOKMARK NOT DEFINED.
4	IMPROVING MT AND SERVICES BEYOND PLUTO	ERROR! BOOKMARK NOT DEFINED.
4.1	TRANSLATION QUALITY	ERROR! BOOKMARK NOT DEFINED.
4.2	TRANSLATION INTERFACE.....	ERROR! BOOKMARK NOT DEFINED.
4.3	MT WEB SERVICES.....	ERROR! BOOKMARK NOT DEFINED.
5	SUMMARY	8

Executive Abstract

The deliverable presents the results of a final survey carried out with the PLuTO working group in WON towards the end of the final year of the project. The aim of the survey was to extend on the original survey published at Month 18 as Deliverable 7.2 and get updated information. However, a slightly different approach was taken in order to get more useful information from the user group. We first provide the motivation behind this approach and introduce the questions posed to the WON members. We then present the answers provided to the survey and discuss what this findings tell us as we turn our focus to the commercialisation of the PLuTO technology beyond the end of the community support.

1 Survey Motivation

The aim of the survey was to extend on the original survey published at Month 18 as Deliverable 7.2 and get updated information. However, we decided to take a slightly different approach in order to extract more valuable information and opinions rather than get further confirmation of facts that were unlikely to have changed, e.g. that the Asian languages were the most important, that machine translation is needed, and the search tools that are used by patent professionals.

Instead of multiple choice or list style questions, we posed a number of questions that required slightly more detailed answers and in some cases opinions and thoughts. This was done as we intended to only survey the smaller working group with WON who have been engaging with the consortium over the course of the project. Our hope was that by engaging them on a more intimate level might encourage them to be more forthcoming with information as opposed to an impersonal survey to the wider user group.

2 Survey Recap

The survey comprised 9 questions in total which were designed to get insight from the perspective of the patent professional into the current and proposed activities stemming from the PLuTO technology. We sought gather a variety of information on items such as pricing, the value of translation technology, and how best to pitch the products. The full set of questions is provided below:

- Q1. PLUTO currently provides translation between English and French, German, Spanish, Portuguese, Chinese, Japanese, and soon Korean and Russian. What other languages or combinations (perhaps not including English) might you consider important?*
- Q2. Aside from the major search tools, we have come across a number of smaller search services and analytic tools for patent information, e.g. Intellixer, S4IP, InfoApps, CrowdIPR, Patentics, Valipat, etc. Do you personally use any of these services and are there others which you are aware of and/or make use of?*
- Q3. Full text machine translation (particularly of Asian patents) is becoming increasingly predominant in the larger patent search tools e.g. PatBase, Total Patent, Orbit. a) Do you find these translations useful? b) Do they reduce your need for on-the-fly translation?*
- Q4. What are your general thoughts on the EPO/Google Patent Translate collaboration and service?*
- Q5. One of the challenges in commercialising our service is choosing the right price and pricing model. If you are willing or able to share, can you provide details on your experiences with the various subscription prices and models (annual, monthly, per seat, per action) you have encountered with services you use?* *any information you provide will be treated with the strictest confidence and will not be shared explicitly with other parties*
- Q6. Another challenge we face is how best to pitch our service to potential users/customers. It is difficult to demonstrate translation quality anecdotally and there are many benefits beyond just that. To that end, in terms of IPTranslator: a) What benefits do you see in IPTranslator in your daily work? (e.g. time saving, discover new patents) b) What might be the most effective way to present this argument in favour of IPTranslator?*
- Q7. Machine Translation can only be used up to a certain point before human translations are needed. What are your thoughts on the possibility of a feature through which a third party human translation could be requested directly through IPTranslator for a given patent?*

Q8. *We are also exploring the development of an online tool for patent search management and collaboration called ParonPro. What are your thoughts on the opportunities and obstacles for such a service?*

Q9. *Finally, as we look to develop the commercial opportunities around our translation technology we are investigating alternative applications within the IP industry, e.g. integration with smaller search/analytic tools, full-text translation for large search vendors, and a productivity aid to professional translators. Do you see any other possible applications within the IP industry?*

3 Survey Responses

We had five respondents in total to the survey who provided a variety of answers which we discuss on the question by question basis below.

Q1. *PLUTO currently provides translation between English and French, German, Spanish, Portuguese, Chinese, Japanese, and soon Korean and Russian. What other languages or combinations (perhaps not including English) might you consider important?*

The consensus here was that with these translation directions accounted for we are covered at the least in the medium term.

Q2. *Aside from the major search tools, we have come across a number of smaller search services and analytic tools for patent information, e.g. Intellixer, S4IP, InfoApps, CrowdIPR, Patentics, Valipat, etc. Do you personally use any of these services and are there others which you are aware of and/or make use of?*

The respondents did not use any of the tools mentioned in their organisations. We suspect that this is because these services are smaller, less comprehensive tools aimed at professionals with less of a tools budget. The WON members all work for large multinational corporations with big budgets to subscribe to the all-encompassing tools such as Thomson Innovation.

Q3. *Full text machine translation (particularly of Asian patents) is becoming increasingly predominant in the larger patent search tools e.g. PatBase, Total Patent, Orbit. a) Do you find these translations useful? b) Do they reduce your need for on-the-fly translation?*

On the whole, these full-text translations are useful and naturally limit the need for on-demand translation when they exist. That being said, quality remains an issue in these translations and it is not always possible to make a relevancy judgment. Another issue that arose was the fact that it was difficult to export or manipulate the full-text translations *in-situ* in the various tools, such as exporting the translation and sharing with colleagues.

Q4. *What are your general thoughts on the EPO/Google Patent Translate collaboration and service?*

Feelings towards the Google Patent Translate service at the EPO were mixed amongst the respondents. Some feel the quality is good while think the quality is generally poor. However, it was interesting to get the sense from the WON users that they appreciate that Google is a large, professional organisation and have the capacity to continue to be a strong

competitor. This is something we will continue to face operating as a small entity in the industry.

Q5. One of the challenges in commercialising our service is choosing the right price and pricing model. If you are willing or able to share, can you provide details on your experiences with the various subscription prices and models (annual, monthly, per seat, per action) you have encountered with services you use?* *any information you provide will be treated with the strictest confidence and will not be shared explicitly with other parties

For this question, the similar sentiments to those we have come across previously were echoed; namely that flat subscription fees are preferred because customers know up front what to budget for. That being said, there is precedent in the industry for different types of models such as per person and pay-a-you-go.

Q6. Another challenge we face is how best to pitch our service to potential users/customers. It is difficult to demonstrate translation quality anecdotally and there are many benefits beyond just that. To that end, in terms of IPTranslator: a) What benefits do you see in IPTranslator in your daily work? (e.g. time saving, discover new patents) b) What might be the most effective way to present this argument in favour of IPTranslator?

Many of the value propositions we already use were echoed by the respondents here, for instance, the ability to understand foreign patents and the efficiency gains to be had using the tool. Other points noted was the usefulness of the translations specifically for geographically targeted freedom to operate searches.

Q7. Machine Translation can only be used up to a certain point before human translations are needed. What are your thoughts on the possibility of a feature through which a third party human translation could be requested directly through IPTranslator for a given patent?

This suggested was roundly welcomed with one respondent pointing out that they only consistently use the same translation for Japanese translation and for all other languages they operate on a case by case basis. A consistent service provided through IPTranslator would appeal to them.

Q8. We are also exploring the development of an online tool for patent search management and collaboration called ParonPro. What are your thoughts on the opportunities and obstacles for such a service?

The respondents believe that this has the potential to be a useful service. Opportunities lie in the fact that it could become an integral part of a workflow if there was buy in from a whole research team. Obstacles would be getting teams to change their current workflows which have been fine tuned (although perhaps are still not as efficient as they could be).

Q9. Finally, as we look to develop the commercial opportunities around our translation technology we are investigating alternative applications within the IP industry, e.g. integration with smaller search/analytic tools, full-text translation for large search vendors, and a productivity aid to professional translators. Do you see any other possible applications within the IP industry?

The respondents not have any suggestions for this question.

4 Summary

The results of this survey have been useful in confirming some of the information we have heard over the course of our engagements with the intellectual property community. It is clear we are going in the right direction in terms of our languages, pricing models, and pitch. It is interesting to note that competition manifests itself not in companies who offer a similar service but rather in an alternative service (bulk translation) that reduces the need for on-demand translation.

This information has feed into our exploitation thought process insofar as we have realised that providing this type of service is also something that we have explore. This is reflected in Deliverable 9.2 in our discussion around providing bulk translation to the patent tool vendor target market.