



**COUNCIL OF
THE EUROPEAN UNION**

Brussels, 3 September 2010

13166/10

**Interinstitutional File:
2010/0198 (CNS)**

PI 96

NOTE

from: General Secretariat of the Council
to: Working Party on Intellectual Property (Patents)

No. prev. doc.: 13076/10 PI 95

Cion prop.: 11805/10 PI 77 + ADD 1 + ADD 2

Subject: Proposal for a Council Regulation (EU) on the translation arrangements for the European Union patent
- Background note on machine translations

Delegations will find in [Annex](#) a background note on machine translations for the EU Patent, drawn up by the Commission services, as a basis for discussions at the meeting of the Working Party on Intellectual Property (Patents) on 8/9 September 2010.

MACHINE TRANSLATIONS FOR THE EU PATENT**BACKGROUND NOTE****DG MARKT****Introduction**

The single European Union patent ("EU patent") must strike the right balance between the objective of cost-effectiveness and the dissemination of patent information and technical knowledge in different EU languages. Striking the right balance between these two objectives has been one of the most difficult issues in the discussions on possible translation arrangements for the EU patent.

In 2008, the Slovenian Presidency¹ explored several possibilities in this regard. A system of machine translations was proposed as a way to reduce translation costs for the users, whilst ensuring the availability of patent documents in all official EU languages. It was proposed that such translations would be carried out by a central service on the request of any interested party, free of charge. Such translations would be available for information purposes only and would have no legal effect. The proposal of such a machine translation system was supported by a large majority of Member States. A number of them, however, underlined the importance of a high level of quality for such translations.

The proposal for a Council Regulation on the translation arrangements for the EU patent² recognises the importance of providing patent information through machine translations. The proposal identifies machine translations as one of the accompanying measures to be established, together with the creation of the EU patent. Discussions on the proposal at the Council Working Parties, of 14 and 28 July 2010, showed that a number of Member States believe that the availability of machine translations is one of the major elements in the overall translation regime for the EU patent. They have requested further clarifications on the current state of play of the programme, as well as an outline of the possible ways forward.

¹ Council documents 6985/08 and 8928/08.

² COM(2010) 350 final, adopted 30/06/2010.

In view of these requests, this background note from Directorate General for Internal Market and Services (DG MARKT) seeks to outline the main objectives, the current state of play, the major challenges and the way forward for the development of the machine translation programme.

I. Objectives

The current rapid development of various machine translation programmes offers a unique possibility both for the users of the patent system and the examiners of patent offices to use this technology in order to facilitate the timely availability of necessary technical information and the search for prior art in multiple languages. However, any generally available machine translation programme must first be customised in order to translate highly technical and complex patent documents with the required level of accuracy.

A specialised machine translation programme for patent documents has already been launched by the Europe Patent Office (EPO) in 2004. The overall objective of this programme is providing an automated translation service of sufficient quality to make the technical content of patent documents understandable to a technically qualified person.

The programme has two dimensions – external and internal. It is first of all intended to foster the dissemination of patent information and technical knowledge to the users of the patent system. Internally, the programme facilitates, speeds up and in some cases enables the process of search for relevant prior art, by the EPO examiners, in multiple languages.

In addition, the programme is also seen as a contribution by the EPO to the efforts of creating a single patent for the territory of the EU. The EPO supports the creation of a simplified and cost-effective EU patent and machine translations are seen as one of the key elements in achieving this. The single EU patent thus creates further incentives for the development of the machine translation programme at the EPO.

The proposal for a Council Regulation on the translation arrangements for the EU patent implies that machine translations for the EU patent would be based on the machine translation programme currently developed by the EPO. It is specified that whereas, currently, machine translations are available in a limited number of languages, the programme would need to be extended to cover all EU official languages. To this end, the Commission should cooperate with and support the efforts of the EPO to roll out the existing programme.

II. State of play

The machine translation programme of the EPO already ensures the availability of customised translations of patent specifications in a limited number of languages. The first language pairs German-English / English-German and Spanish-English / English-Spanish became available in 2008. Since 2009, French-English / English-French and Italian-English / English-Italian are also available but are still being improved. The existing language pairs already receive a significant number of machine translation requests. On the basis of the data collected in the first quarter of 2010, the total number of requests for machine translations in all language pairs should reach more than 1 million requests per year. Currently work is under way to provide subsequent language pairs for Swedish-English / English-Swedish and Portuguese-English / English-Portuguese. Further work is also planned for Dutch, Finnish, Greek and Romanian.³

On 26 August 2010, a document on an extended comprehensive language technology services co-operation programme was submitted for a decision by EPO Administrative Council in October 2010.⁴ If approved, the programme would be launched on 1 November 2010, for the duration of four years. In particular, this foresees the acceleration of the development of the language technology services, including a viable machine translation programme as envisaged in the Commission proposal on the translation arrangements for the EU patent. The document emphasizes that, based on previous experience, future work will be a collaborative exercise between various players. The EPO will focus on areas where it can add value and will collaborate with partners both from the public and private sectors (for example, national patent offices, IT service providers) in their respective areas of expertise.

³ Details of the languages where machine translation is available or is planned is available in the EPO Status Report on the European Machine Translation Programme document CA/T 6/09.

⁴ EPO document on Cooperation programme on European language technology services for patents CA/29/10

The programme will consist of four phases as follows:

1) *Building a corpora of patent documents* – building a set of patent documents to be used for the "training" of translation engines, in order to provide customised translations of patent documents.

2) *Language technology services delivery* – the provision of electronic translation services across all EPO languages into and from English, French and German. This will involve launching a call for tender for the translation services. This phase of the project relies on the build up of sufficient patent corpora across all EU languages, and a suitable provider of language services for all EU languages. Where this is not possible, manual translations to increase the number of patent documents in a respective language may be required.

3) *Integration of language technology services* from phase 2 into patent information tools and services. These would be of interest to both patent examiners (at the EPO and national patent offices) and the general public.

4) *Maintenance* – keeping the tools updated and functioning properly.

The EPO is continuing and extending its work in developing and improving the machine translation programme, with a view of contributing to the efforts of establishing a fully fledged programme for the future EU patent.

The European Commission has a keen interest in developing machine translation programmes. In 2009, an inter-service task force recommended the development of a new machine translation system, as a Commission service (MT@EC) in all EU languages. The technical implementation of the baseline system will last from October 2010 to March 2011 and will be funded by the Commission as part of the ISA programme.⁵ An elaboration and development phase will follow, to be finalised by the end of 2012.

⁵ Interoperability solutions for European public administrations – under the responsibility of Directorate General for Informatics (DG DIGIT).

The Commission has financially contributed to various projects supporting the development of machine translation systems. Currently Directorate-General for Information Society and Media (DG INFSO) oversees several projects relevant for the machine translation programme within the 7th Framework Programme (FP7) and the ICT Policy Support Programme (CIP). Both programmes involve the EU contributing 50% (current CIP projects) or 75% (FP7 projects) of funding for the duration of the projects.

Relevant 'live' projects include:

- PLuTO (Patent Language Translations Online) which specifically targets patent online machine translations. This runs from 1 April 2010 to 31 March 2013. Its total cost is 4.36 million EUR.
- LetsMT! which builds an online collaborative platform for data sharing and contributes to machine translations, especially for minority languages. This runs from 1 March 2010 to 31 August 2012. The total cost is 3.34 million EUR.
- iTranslate4 seeks to integrate the best machine translation services of all major European machine translation providers in a single website. This runs from 1 March 2010 to 29 February 2012. Its total cost is 3.94 million EUR.

In September 2010, DG INFSO will also publish two new calls for tender for projects related to the development of machine translations.⁶

III. Challenges

To be available as an accompanying measure for the creation of the EU patent, the current EPO machine translation programme faces two major challenges:

- extending the programme to all official languages of the EU; and
- improving the quality of all translations in all language pairs.

⁶ Namely, the 7th call for proposals under the language technologies objective in the ICT Work Programme 2011-2012T and the SME-DCL (SME digital content online) will be launched. The EU will fund 75% of the costs for selected projects with the budget of 85 million EUR.

The extension of the programme to all official languages of the EU requires significant efforts from all the parties concerned. Currently, seven EU official languages are part of the EPO machine translation programme, some are still in the early stages of development. Ways to extend the programme and to roll it out to all other languages are currently being explored.

However, an even more demanding challenge is achieving the necessary accuracy of machine translations for highly complex technical documents. The goal is to allow a technically skilled person to understand the essential elements of the translated patent document.

In case of the currently available language pairs, significant improvements in quality have been achieved. Two years after they first became available, the German-English / English-German and Spanish-English / English-Spanish are now considered as being fully implemented and satisfying the internal quality standard.⁷ This demonstrates, in particular, that the general machine translation engines, that may be quite easily acquired, must be extensively "trained" in order to be successfully used for translations of patents. The current process of developing the engines at the EPO includes building large sets of patent documents and feeding them into the engine to suitably "train" it for the translation of technically complex and precise patent documents. This method requires, in particular, large amounts of patent documents, mainly aligned patent document pairs but also monolingual documents. Currently, the collection of necessary documents is jointly funded by the EPO, in the framework of the co-operation budget, and the national patent offices.⁸ The collaboration with national patent offices is essential for this project. The launch of a renewed cooperation programme on European language technology services for patents is to be approved by the Administrative Council in October 2010. This provides for extended collaboration with national patent offices and external service providers.

⁷ See EPO Status Report on the European Machine Translation Programme document CA/T 6/09.

⁸ EPO document on Cooperation programme on European language technology services for patents CA/28/10, to be replaced by CA/29/10.

IV. Way ahead

The EPO has launched and continues to develop a comprehensive and specialised machine translation programme. The proposed extended cooperation programme on European language technology services for patents indicates the readiness of the EPO to accelerate the development of machine translations.

The Commission services will discuss with the EPO on the type of assistance and support which could be provided to ensure the timely and complete roll out of the machine translation programme. In this process, their priority is ensuring the development of machine translations into all EU official languages and the high quality of these translations. Without prejudice to the substance of the forthcoming discussions with and within the EPO, the Commission services will provide more background on this at the meeting of the Council Working Party on 8-9 September 2010.

The Commission services will report by the end of October on the results of the discussions with the EPO and the envisaged cooperation with a view to providing a timely and fully functioning machine translation system for the EU patent.

In this context, the Commission services will also take due account of the outcome of the discussions which will take place in the October meeting of the Administrative Council of the EPO, on the proposed cooperation programme on European language technology services for patents.
