

EPO's Machine Translation Services Status Update

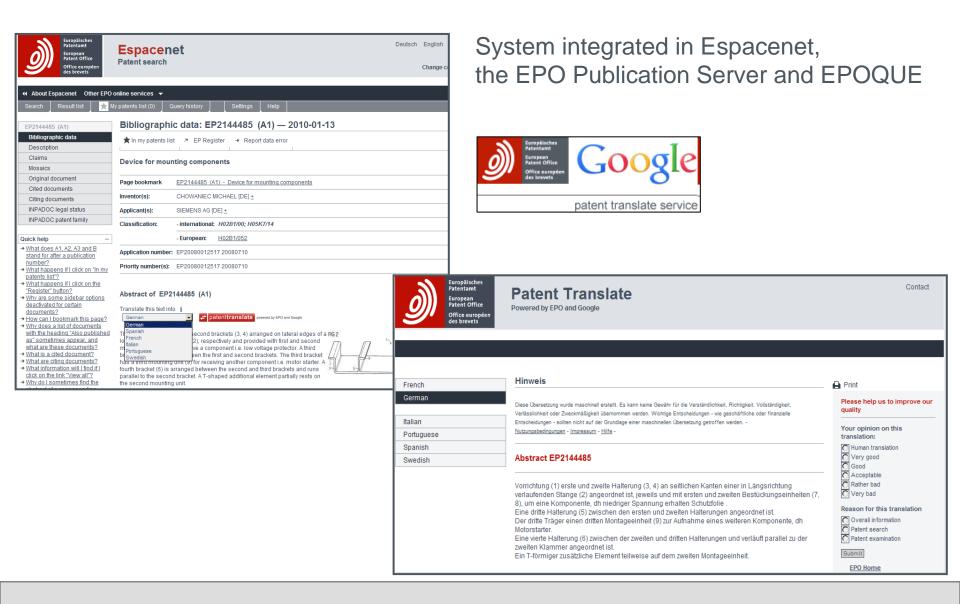
Raimund Lutz Vice-President DG5 Legal and International Affairs

Presentation to IP5 Heads / Industry 5 June 2012





Patent Translate launched on 29 February 2012





Launch of Patent Translate

- Collaboration EPO-Google
- Patent documents and their translation/corresponding documents are prepared and stored in a corpora repository.
- Translation system is trained using this corpora.
- Translation quality check to decide whether the translation is usable for its purpose. Languages with a sufficient quality are offered to the public.



Current achievements

- Corpora Repository contains corresponding patent documents for the following language pairs which was used for translation system training:
 - Batch1: EN-(FR,DE,PT,IT,ES,SV) in production
- From all over the world 35 000 requests per day for Patent Translate
- 90 million different machine translations of complete patent documents can be accessed 'on the fly', using the current language pairs offered.



Next steps

- Additional corpora in order to improve the translation quality for already publicly available languages and to achieve a sufficient quality level for the so far not publicly available languages.
- Next training session includes 7 additional languages including a first corpora batch for Chinese English.
- In 2014 all languages (32 in total including JP/CN/KR/RU) can be expected to be publicly available.



Patent Translate: illustrative example (Description)



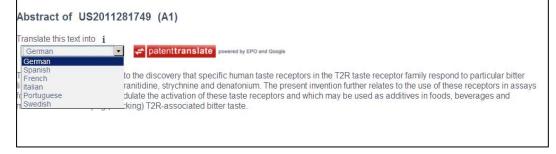


Test it and watch the film











Thank You!

Raimund Lutz Vice-President DG5 Legal and International Affairs