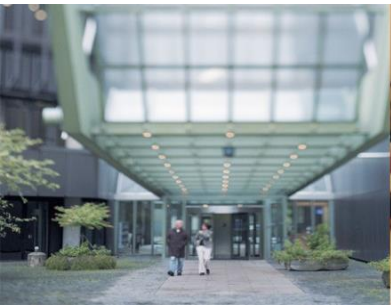


# Machine Translation at the EPO

## Removing language barriers from patent documentation

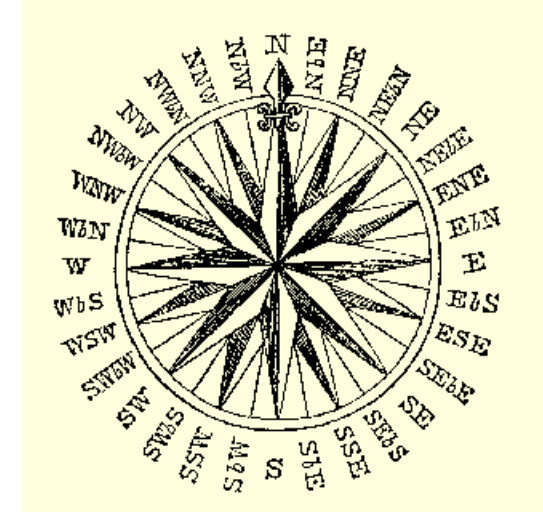
**Paul Schwander**  
**European Patent Office**

**The 5th Workshop on Patent Translation, MT Summit 2013**



# Roadmap

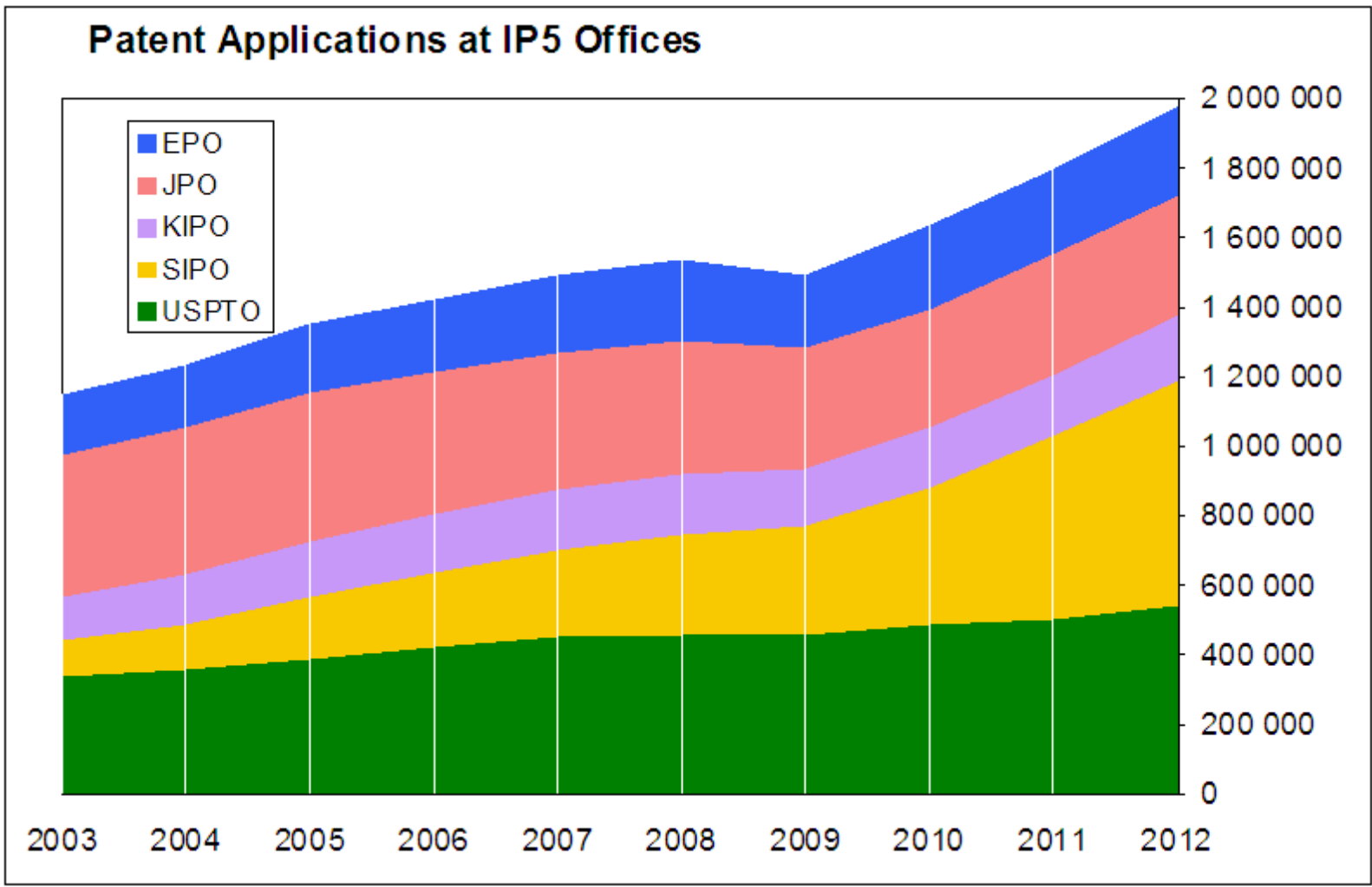
- The context: why is MT strategic?
- Machine Translation @ the EPO:  
state of play and future plans



# Why Machine Translation?

- Reducing the language barrier in the European context: Unitary Patent system supported by MT.
- Access to global patent information for prior art searches.





**Global patent filings rising continuously, especially Chinese applications**  
**IP5= Europe, USA, Japan, China and Korea**



# Accessing Asian languages patent for EPO examiners

CN1201256C C 20050511

Automatically Translated by the Intellectual Property Publishing House of SIPO

Registered system and method of the domain name of native language

Technical field

This invention generally involves the communication network. Particularly, it is not a method and apparatus of naming system of English language that this invention involves a adaptation, addressing and visit the communication network, especially the entity of the WWW directly.

Technological background

Internet, especially is offering a lot of information resources in the WWW , every resource offers some useful information to vast computer network user, goods and service, its typical form is that hyperlink labels the language (HTML) page (webpage). As everyone knows, user use install in they net browser in the computer, such as NAVIGATOR of American Netscape, or INTERNET EXPLORER of American Microsoft, in order to visit their required webpage.


In order to visit a particular webpage, users' net browser must know in the network, namely the address of required targeted website in Internet. Especially in the WWW, its address means the IP address . What the IP address is used is digital format, such as 123.456.78.9. And the website server that every one is regarded as the WWW has a specialized IP address.

However, as webpage quantity in the WWW runs up, design a more direct site selection method, it has used the alias that convenient memorial letter and figure form. According to the site selection system designed newly, usually, the address means a common resource indicating device (URL), its typical grammar is: Protocol (agreement): //WWW . Domain-name (domain name). Domain (land). To WWW, its agreement refers to the hypertext transfer protocol (HTTP). The land has pointed out which Network Entity NE more advanced classification this address belongs to, for example, the land can be com, org, net, or similar name (besides U.S.A., it was this country that showed in other countries in the land, it is the commercial land of Britain as com.uk represented). The domain name makes up the alias of the real IP address with land word separator.

And the domain name can be made up with many stature names, for instance: Sub name 1. Sub name 2 ....Son famous n.. For example, one URL, <http://www.xyz.com> Represent the server in the WWW as the host computer of the websites of xyz Company.

Keep the global address book with special server which is called the name server (DNS) put in Internet, it can map if the domain name of " xyz.com " to regarding as the physical address of the server of the network station host of xyz Company correspondingly, such as 123.456.78.9. The name server has more than one computer separating physically, and carry

Bibl Bib 1/1
[19] 中华人民共和国知识产权局
[51] Int. Cl<sup>7</sup>  
G06F 17/60



## [12] 发明专利说明书

[21] ZL 专利号 01802492.0

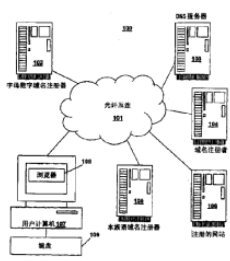
[45] 授权公告日 2005 年 5 月 11 日
[71] 授权公告号 CN 1201256C

<p>[22] 申请日 2001.8.17 [21] 申请号 01802492.0</p> <p>[30] 优先权 [32] 2000. 8. 21 [33] US [31] 09/642,471</p> <p>[86] 国际申请 PCT/US2001/041785 2001. 8. 17</p> <p>[87] 国际公布 WO2002/017204 英 2002.2.28</p> <p>[85] 进入国家阶段日期 2002.4.22</p> <p>[71] 专利权人 金基锡 地址 美国马里兰州 共同专利权人 柳志烈</p> <p>[72] 发明人 金基锡 柳志烈 审查员 蔡 萍</p>	<p>[74] 专利代理机构 中科专利商标代理有限公司 代理人 戎志敏</p>
---	---

权利要求书 2 页 说明书 11 页 附图 5 页

[54] 发明名称 本族语域名的注册系统和方法

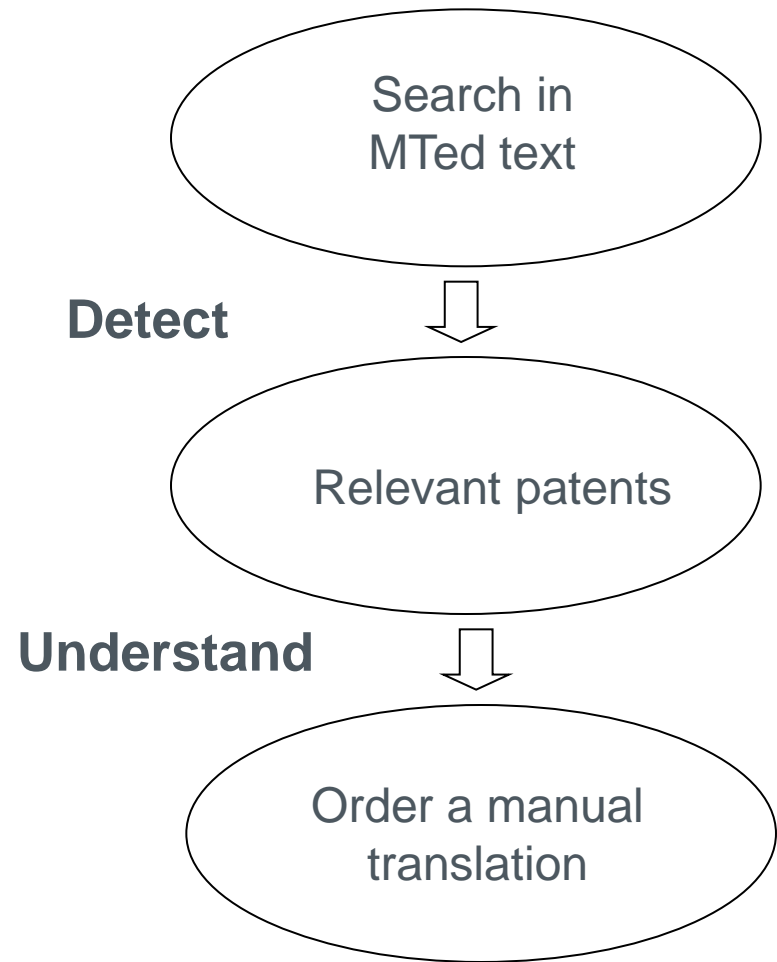
[57] 摘要  
一种域名系统, 包括: 本族语域名注册器 (105), 它用来接收本族语域名, 这个本族语域名包括至少一个非字母数字字符; 申请注册并将本族语域名转换成一个字母数字域名。字母数字域名被用在域名服务器 (DMS) (103) 的域名/IP 地址入口中。



ISSN 1008-4274

知识产权出版社出版

- MT Full-text acquired, ca. 5 million documents.
- An on demand manual translation service offered to examiner
- 5 million patents: manual translation -> 1 day a patent -> 22 years of work for a team of 1000 translators.





## Patent Translate in Espacenet

1. Find a patent document in Espacenet.
2. Choose the part you want to translate - abstract, description or claims.
3. Select the target language, and click the Patent Translate button.

### Description of WO2011131922 (A2)

Translate this text into **i**

German



patenttranslate

powered by EPO and Google

4. Once you see the translation, you can mouse-over to see the original text sentence by sentence.

SYSTEM UND VERFAHREN Für eine Fläche von DMS

BESCHREIBUNG

HINTERGRUND DER ERFINDUNG

Die Erfindung betrifft ein System und Verfahren für ein DMS-Oberfläche, die eine intelligente intelligente Oberfläche DMS und welche einen direkten Anwendung der Turbine Strukturen und

Windturb

Die Erfin

und Flug

Insbeson

Eigensch

DMS bie

The invention relates to a system and method for a surface strain gauge, which can provide an intelligent smart surface strain gauge and which has direct application to wind turbine structures and wind turbine blades and wind turbine struts.

omungsprofil Tragflächen

her intelligenten Oberfläche

Gemäß der Erfindung kann das Material in die Oberfläche einer Struktur in einem Array integriert werden, umfassend getrennte Blätter aus kohlefaserverstärktem Epoxy, die zusammen verbunden sind und denen jeweils eine separat adressierbare Element wie beispielsweise ein Mikrocontroller.

5. Submit your feedback.



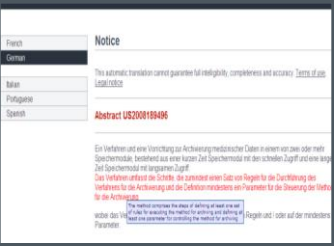


# Patent Translate: how does it work?

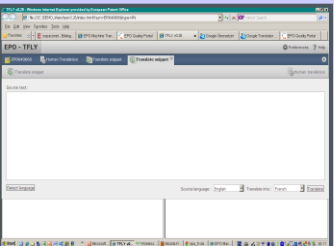
- Result of a collaboration between the EPO and Google
- Patent data represent a huge source of corpora.
- Patent documents and their translation/corresponding documents are prepared and stored in a corpora repository.
- Translation system is trained using this corpora.
- Translation quality assessed before launch: test fit for purpose level.

# Patent Translate: Architecture

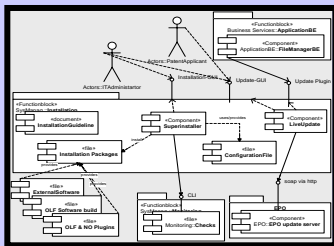
### Espacenet




### Publication Server



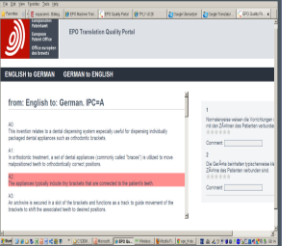
### Bulk

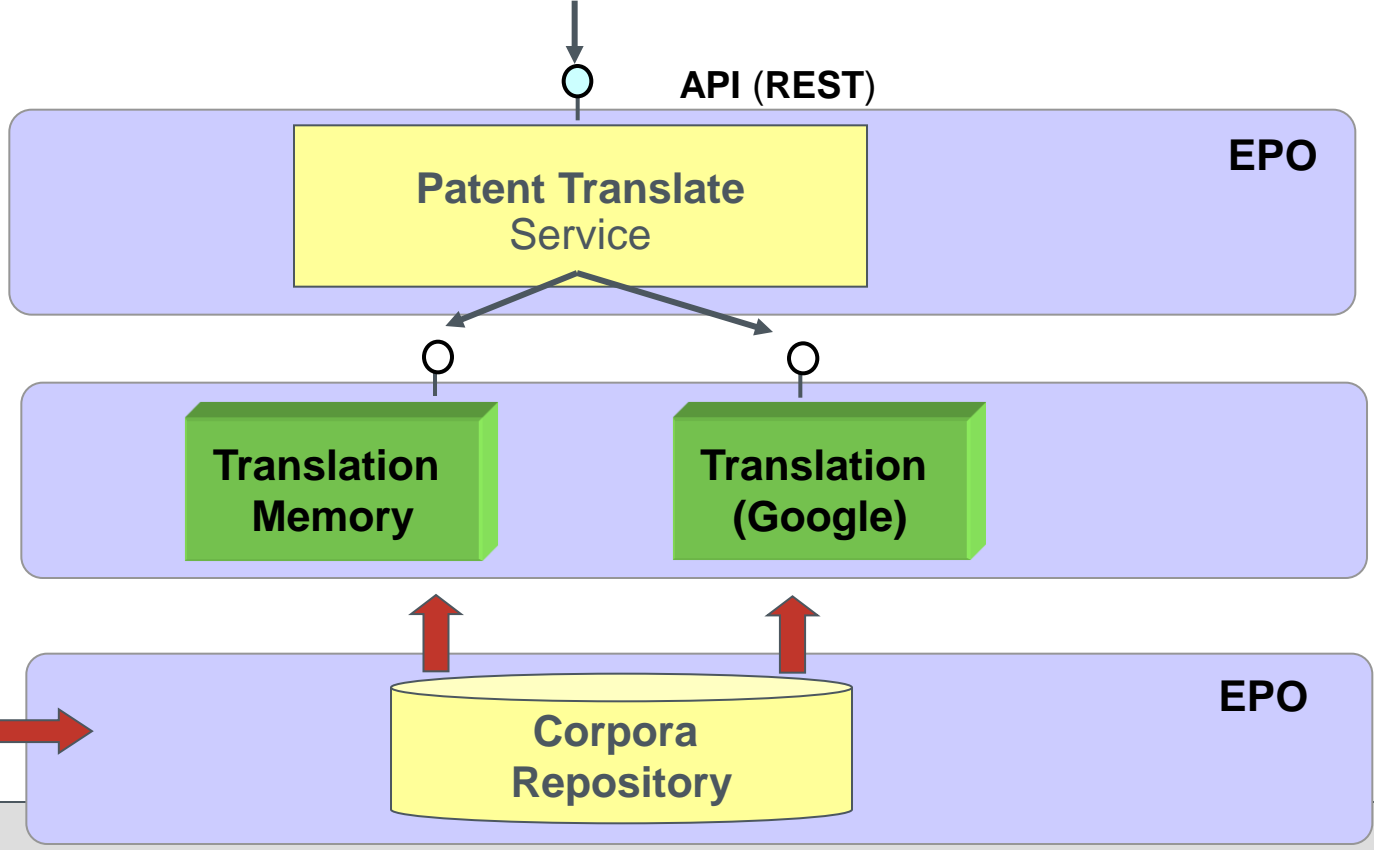


### Examiner tools



### GPI





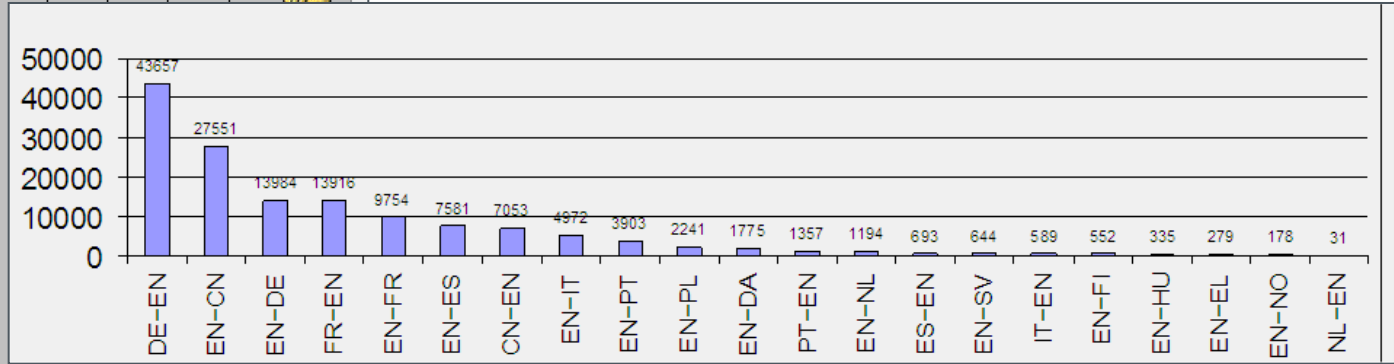
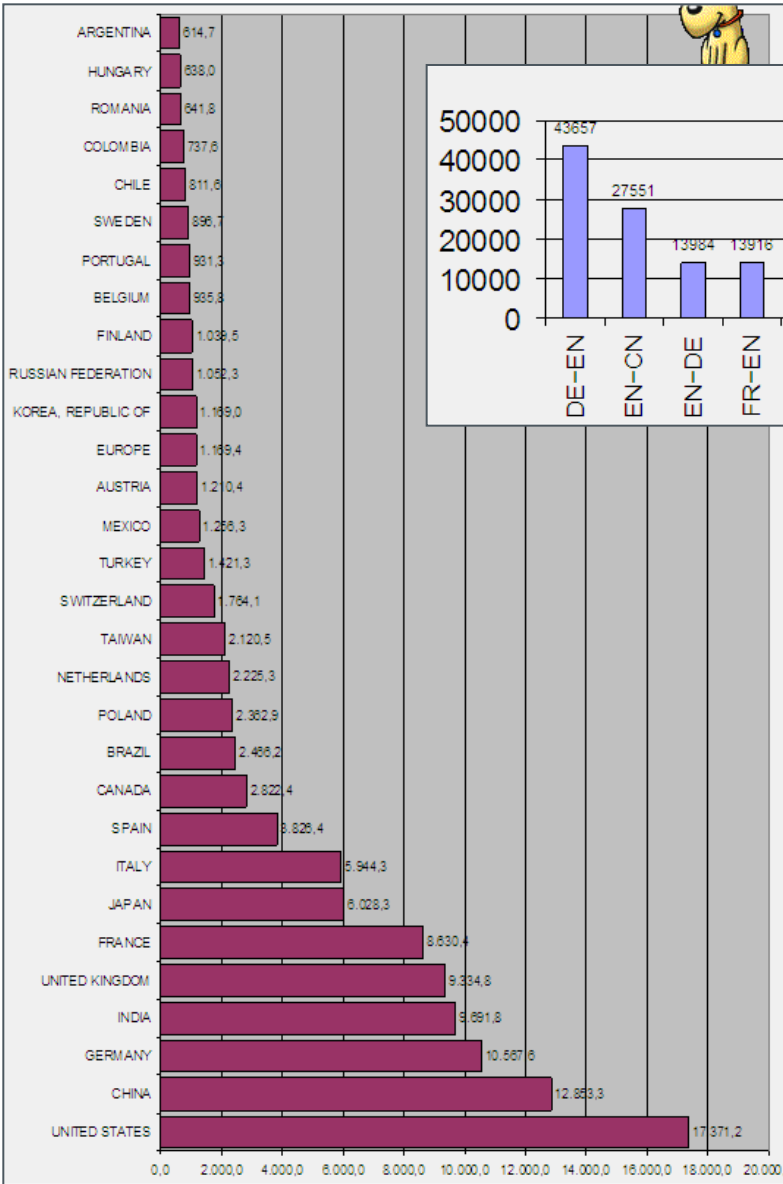
# Quality level ranking

Assessment		Usable for PATENT public	Usable for PATENT search	Usable for PATENT examine
<b>5</b>	Accurate + consistence IPC vocabulary	Yes	Yes	Yes
<b>4</b>	Fluent - consistence IPC vocabulary	Yes	Yes	Yes/No
<b>3</b>	Actionable	Yes	Yes	-
<b>2</b>	May be actionable	Yes/No	-	-
<b>1</b>	Not useful	-	-	-

## Current achievements

- Patent Translate now covers translations between English and 21 other languages: Bulgarian, Chinese, Czech, Danish, Dutch, Finnish, French, German, Greek, Hungarian, Icelandic, Italian, Japanese, Norwegian, Polish, Portuguese, Romanian, Slovak, Slovenian, Spanish and Swedish
- **305 million** different machine translations of complete patent documents can be accessed 'on the fly', using the current language pairs offered = 1500 years of work for 1000 translators if done manually

# Patent Translate usage



# Plans

- Project to be completed end of **2014**:  
32 languages European and Asian.

- **2013-2014**

Turkish, Estonian, Croatian, Latvian,  
Lithuanian, Albanian, Macedonian, Serbian,  
Russian and Korean

# Patent Translate : illustrative example (Description)

- US2006000930 (A1)
- Bibliographic data
- Description**
- Claims
- Mosaics
- Original document
- Cited documents
- Citing documents
- INPADOC legal status
- INPADOC patent family

**Description: US2006000930 (A1) — 2006-01-05**

 In my patents list  
  EP Register  
  Report data error

 Print

## Servo valve for controlling an internal combustion engine fuel injector

### Description of US2006000930 (A1)

Translate this text into


 powered by EPO and Google

The EPO does not accept any responsibility for the accuracy of data and information originating from other authorities than the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

### Description of US2006000930

- French
- German
- Italian**
- Portuguese
- Spanish
- Swedish

Translate this text into

- Italian
- German
- Spanish
- French
- Italian**
- Portuguese
- Swedish

#### Avviso

La presente traduzione è stata eseguita mediante sistema meccanizzato. Pertanto non si garantisce la sua intelligibilità, precisione, completezza, affidabilità o idoneità a impieghi specifici e se ne sconsiglia l'uso in sede di decisioni di natura commerciale o finanziaria. - [Condizioni di utilizzo](#) - [Avviso a norma di legge](#) - [Aiuto](#) -

#### Description US2006000930

[0001] La presente invenzione si riferisce ad una servovalvola di controllo di un motore a combustione interna iniettore di combustibile.

[0002] Come è noto, un iniettore comprende un corpo iniettore che definisce un ugello per iniettare combustibile nel motore, e ospita un'asta di comando mobile lungo un rispettivo asse di attivare una spina di chiusura dell'ugello.

Il corpo iniettore ospita anche una valvola elettromagnetica servo controllo comprendente una camera di controllo delimitata assialmente da un lato dall'asta di comando, e dall'altra da una parete di fondo avente un foro di uscita che, all'esterno della camera di controllo, esce assialmente all'interno di una sede conica.

La servovalvola di controllo comprende inoltre un otturatore che a sua volta comprende una sfera che impegna la sede conica, ed è attivato da un elettromagnete per spostarsi assialmente da e verso il sedile per aprire e chiudere il controllo.

Più specificamente, l'otturatore esercita la pressione del carburante nel foro di uscita della spinta assiale di una molla e l'elettromagnete non è eccitato.

[0003] domanda di mercato attuale.

[0004] Quando sono sottoposti a tensione, tuttavia, attuatori piezoelettrici possono esercitare spinta ma non tirare, e pertanto non può essere usato nelle soluzioni note sopra descritte.

[0005] Inoltre, attuatori piezoelettrici produrre uno spostamento relativamente piccolo, così che, per raggiungere la necessaria di carburante tratti di flusso di scarico, sistemi di amplificazione viaggio devono essere forniti, o parzialmente il foro di uscita zona di tenuta aumentata.

Da un lato, sistemi di amplificazione di viaggio sono indesiderabili, principalmente essendo complessi ed ingombranti, e, dall'altro, un aumento nella zona di sigillatura aumenterebbe la forza assiale esercitata dalla pressione del combustibile sull'otturatore nella posizione chiusa, in modo che il precarico della molla dovrebbe essere aumentato per mantenere l'otturatore chiuso, e una maggiore

The control servo valve also comprises a shutter which in turn comprises a ball engaging the conical seat, and is activated by an electromagnet to move axially to and from the seat to open and close the outlet hole and so vary the pressure inside the control chamber.

[0002] As is known, an injector comprises a control chamber bounded outside the control chamber, comprising a control rod extending axially through the control chamber, and a ball engaging the conical seat, and so vary the pressure inside the control chamber by varying the pressure of the fuel in the outlet hole. When the pressure is kept the outlet hole closed when the pressure is increased.

[0003] Current market demand is for a simple and reliable solution.

[0004] When subjected to voltage, piezoelectric actuators can exert force but not pull, and therefore cannot be used in the known solutions described above.

the EPO; in particular, the EPO does not guarantee that they are complete, up-to-date or fit for specific purposes.

uses a control rod extending axially through the control servo valve outlet hole which, in turn comprises a ball engaging the conical seat, and so vary the outlet hole and so vary the pressure inside the control chamber.

used in the known solutions described above.



# MT Example

14.

Dispositif selon la revendication 13, comprenant un dispositif de commutation permettant d'activer ledit dispositif lors de l'insertion de ladite clé dans une serrure.

15.

Disp A device as claimed in claim 13 including a switch arrangement used to activate said device upon insertion of said key in a lockset. age comprend une indication visuelle de l'état de la serrure comme mis à jour par le dispositif de traitement.

# MT Example

1.  
Ce qui est revendiqué est la suivante: 1.  
Procédé de synchronisation d'une représentation d'affichage de données avec des changements associés aux données, le procédé comprenant:  
déterminer un changement de données, les données associées à une représentation graphique des données;  
calculer une valeur de temps préféré pour la synchronisation de la représentation graphique de la variation déterminée de données, la valeur de temps préféré basé au moins partiellement sur une durée de temps nécessaire pour effectuer au moins une synchronisation préalable de la représentation graphique;  
la comparaison d'une valeur de temps actuelle à la valeur de temps préféré, et  
**la détermination d'un temps de synchronisation sur la base de résultats de la comparaison, le temps de synchronisation identifiant un temps d'initier une synchronisation de la représentation graphique.**
2.  
Procédé selon la revendication 1, dans lequel les données comprennent des attributs d'objets dans une banque de propriétés.  

determining a synchronization time based on results of the comparison, the synchronization time identifying a time to initiate a synchronization of the graphical representation.
2.  
Procédé selon la revendication 1, dans lequel les données comprennent des attributs d'objets dans une banque de propriétés.

## Topics around Patent MT

- CLIR or not?
- OCR combined with MT for non-digitised collections
- Fit for purpose quality assessment
- Perception of the quality



# Conclusion

MT is more than ever a must in the context of the global patent documentation:

- The size of patent collections to search is increasing and systematic manual translation is not an option.
- MT has proven to be fit for purpose
- Quality will continue improving



# Thank You

[www.epo.org](http://www.epo.org)

[pschwander@epo.org](mailto:pschwander@epo.org)