

Application of machine translation system in JPO

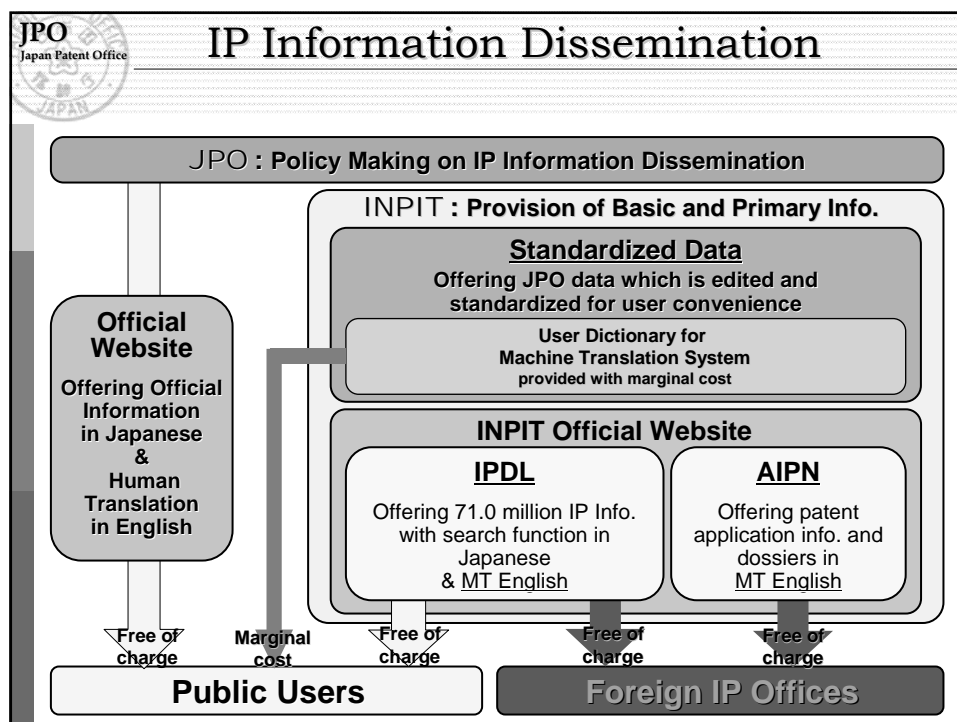
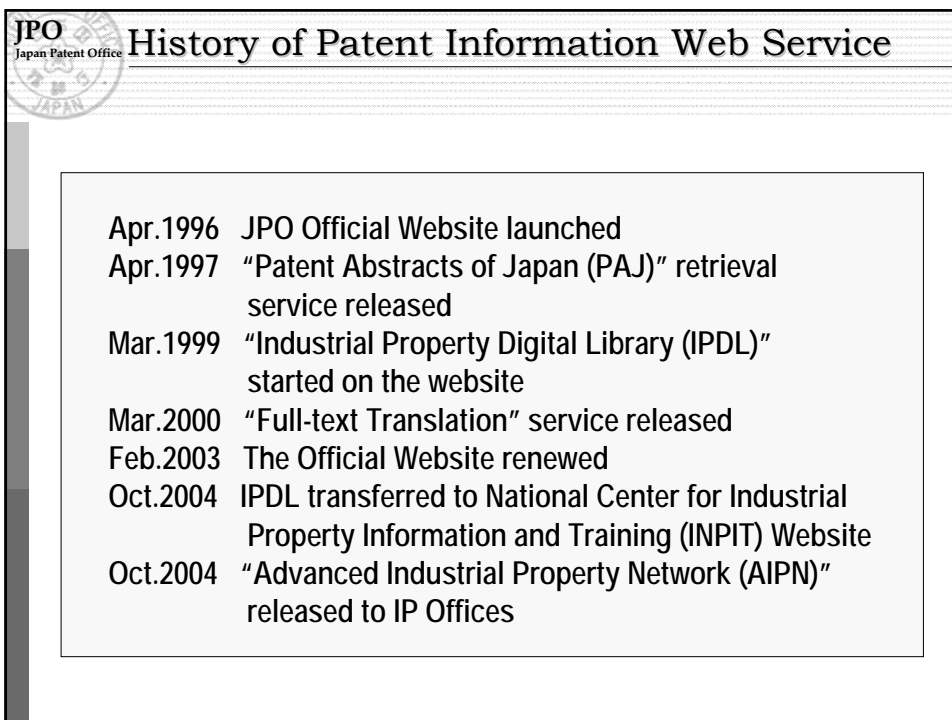
August 29, 2009

Japan Patent Office
Susumu BANI



Contents

- History of Patent Information Web Service
- IP Information Dissemination
- IPDL (Industrial Property Digital Library)
- AIPN (Advanced Industrial Property Network)
- Approaches for Improved Translation Result
- Optimization Plan
- Future Image of Machine Translation
- IP5 Foundation Project



JPO
Japan Patent Office

IPDL (Industrial Property Digital Library)

The screenshot shows the IPDL website with several callouts pointing to specific database categories:

- Patent Databases with MT service:** Points to the 'Patent & Utility Model' section.
- Design Database with MT service:** Points to the 'Design' section.
- Trademark Databases:** Points to the 'Trademark' section.

English Databases on IPDL

- 1. Patent Databases**
 - DB with Machine Translation Service**
 - Patent & Utility Model Gazette DB
 - Patent & Utility Model Concordance
 - F/I/F-term Search
 - Human Translation Database**
 - Patent Abstracts of Japan (PAJ)
- 2. Design Database**
 - DB with Machine Translation Service**
 - Design Gazette DB
- 3. Trademark Database**
 - Human Translation Database**
 - Japanese Trademark Database
 - Japanese Figure Trademarks
 - Japanese Well-Known Trademark
 - List of Goods and Services

JPO
Japan Patent Office

IPDL (Industrial Property Digital Library)

Full-text Patent Publication by MT

Full-text Patent Publication

- All information (except drawings) is translated by machine translation system.

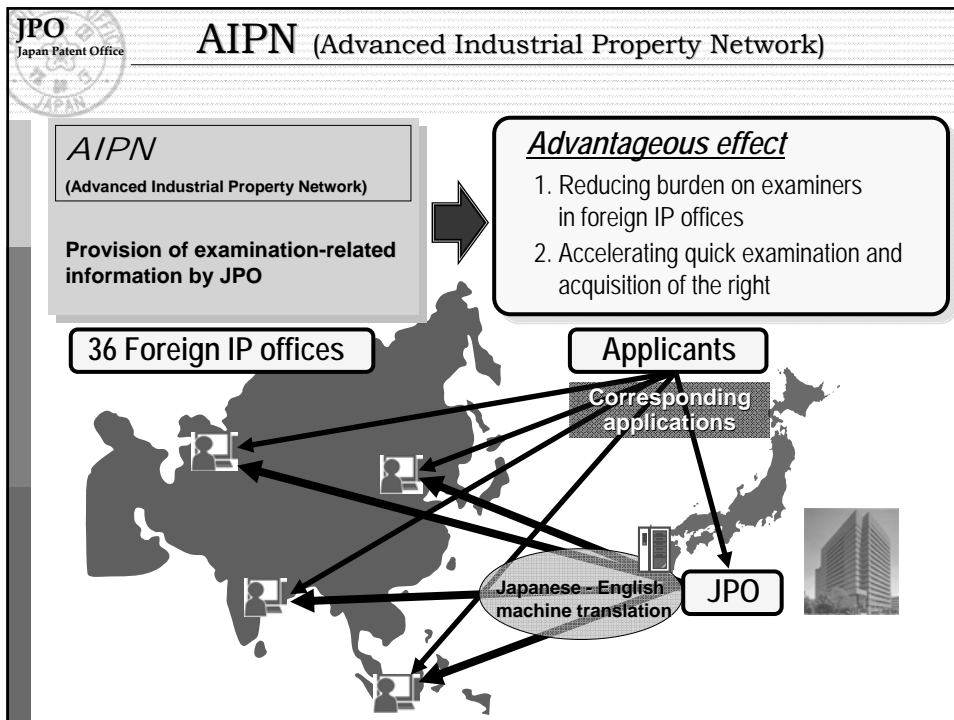
The screenshot shows a patent document interface with two callouts:

- Full-text by MT:** Points to the 'CLAIMS' section, which contains the following text:

CLAIMS

[Claim(s)]
[Claim 1]

In a photovoltaic power generation apparatus provided with a solar cell array which carried out series parallel connection of two or more solar cell modules, and formed them, and a power conditioner which has a converter which carries out conversion into ac of the direct-current generated output from this solar cell array,
An actinometer is attached so that said solar cell array, an installation direction, and installation ***** may become the same, and a temperature sensor is
- Selected Drawing:** Points to a technical drawing of a solar cell array circuit diagram.



JPO
Japan Patent Office

AIPN (Advanced Industrial Property Network)

- Japanese examination-related information* is available for inspection in English by means of Japanese-English machine translation
- Available for foreign IP offices
- Free of charge

*Examination-related information

- Status of the application or prosecution history
- Official gazettes
- Cited documents/patent family information
- File wrapper documents (Dossier)



AIPN (Advanced Industrial Property Network)

File Wrapper Document List

- Examination-related information is available in English through machine translation system.

Application information with legal status

Dossier List
Office actions
Search reports
Written argument
Examiner's note
etc.

Patent application number 2001-123456

File Wrapper Document List - Click document title to display contents.
The dates in the list are shown in the format of "dd.mm.yyyy"

Application data and Transaction history

- 20.04.2001: Request for a Patent
- 20.04.2001: Description
- 20.04.2001: Drawings
- 20.04.2001: Abstract
- 02.10.2001: Ex Officio Correction of Defect of Abstract
- 23.04.2002: Request for Examination (Others' Request)
- : --- : --- Certified or Appended Information
- 28.05.2002: Written Notice (others' request)
- 22.07.2003: Search Report by Designated Searching Authority
- 12.08.2003: Assessment on Search Report by Designated Searching Authority
- 02.09.2003: Notification of Reasons for Refusal
- 04.11.2003: Written Argument
- 04.11.2003: Written Amendment
- 21.04.2004: Decision to Grant a Patent

[Translation done.]



AIPN (Advanced Industrial Property Network)

- Examination-related information is available in English.
- All information is translated by machine translation system.

ENGLISH JAPANESE

Note: Japanese environment is required to properly display Japanese characters. You must install and use a TIFF image plug-in on your system in order to view image files directly.

Disclaimer
This English translation is produced by machine translation and may contain errors. The JPO, the INPI, and those who drafted this document in the original language are not responsible for the result of the translation.

Notes:
1. Untranslatable words are reflected with asterisks (****).
2. Texts in the figures are not translated and shown as it is.

Translated: 16:22:55 JST 06/05/2008
Dictionary: Last updated 05/20/2008 / Priority:

[Document Name] Description

[Title of the Invention] Flexible copper-clad sheet

[Claim(s)]

[Claim 1] In the flexible copper-clad sheet with which the copper layer was formed on the flexible polymer base material (1) The surface of a flexible polymer base material is mostly dotted with the independent minute metal membrane at homogeneity. (2) The part which is not dotted with the metal membrane with the minute surface of a flexible polymer base material has average depth 0.001-2.0micrometer impression structure from the surface, and covers a minute metal membrane and impression structure on the surface of (3) flexibility polymer base material. The flexible copper-clad sheet characterized by forming the intermediate metal layer and the copper layer in this order.

ENGLISH JAPANESE

Note: Japanese environment is required to properly display Japanese characters. You must install and use a TIFF image plug-in on your system in order to view image files directly.

used by machine translation and may contain errors. The JPO, the INPI, and those who drafted this document in the original language are not responsible for the result of the translation.

located with asterisks (****), translated and shown as it is.
2008 / Priority:

Notification of Reasons for Refusal

Application for patent 2001-123456
15(2003) August 12
M, Nobuhiro 9341 3S00
nt NISHIKAWA, Shigeaki
nt Law Section 29(2)

be refused for the reason mentioned below. If the applicant has any argument an argument should be submitted within 60 days from the date on which this had.

Reason

In the each claim listed below of this patent application should not be granted a ion of Patent Law Section 29 (2) for the reason that the claimed invention(s) made by persons who have common knowledge in the technical field to which the

Notification of Reason for Refusal

Description Document



AIPN (Advanced Industrial Property Network)

Patent Examination Cooperation

Patent Prosecution Highway (PPH)

On condition that a patent application has been determined to be allowable in the first office, the application is qualified for accelerated examination through a simple procedure according to a request from an applicant in the second office.

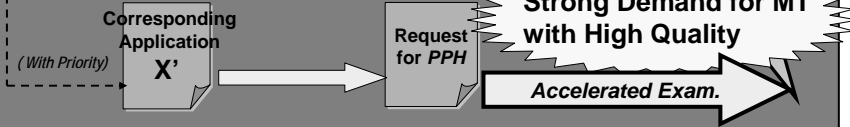
- It assists applicants to obtain a foreign patent expeditiously.
- It reduces burden of examination and enhances quality of examination for Offices.

First Office



Second Offices utilize the search and examination result of First Office, which is translated by MT.

Second Office



Approaches for Improved Translation Result

Pre-modification

Pre-modification for machine translation

Pre-modification is necessary to reduce mistranslation arise from format error.

examples

1. Replacement (elimination) of tag information

Unnecessary line feed: ¥r¥n¥r¥n¥r¥n → ¥r¥n

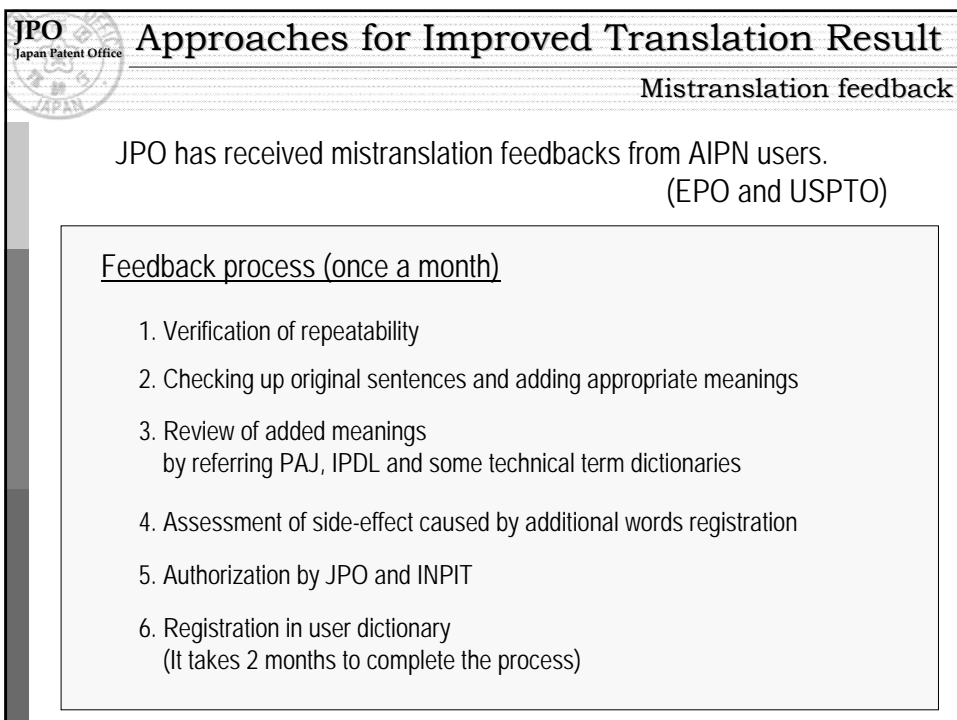
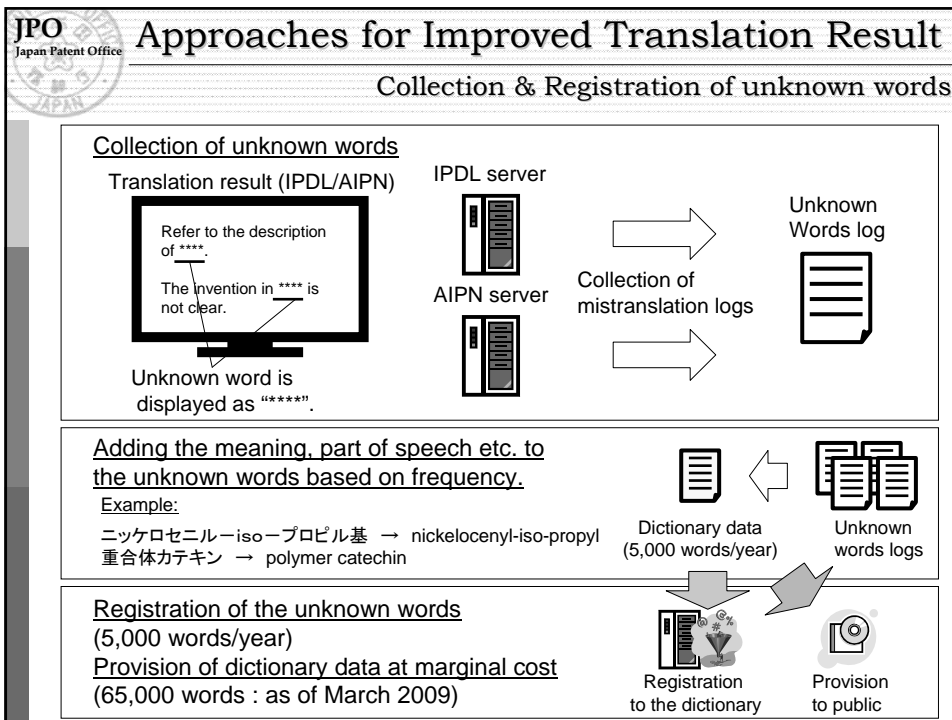
2. Deleting unnecessary spaces

「引用文献等一覧」 → 「引用文献等一覧」
"The list of cited documents"

3. Pre-formatting

Document number: "yyyyxxxxx" → "yyy-xxxxx"

Avoiding Consequent quotation: "bcd(kyz)" → "bcd(kyz)"





Approaches for Improved Translation Result

Research for accurate translation result

Brief summary of the research

Numbering

Some types of numbering cause mistranslation.

example

| | |
|---|--|
| Original sentence | この出願は、特許請求の範囲の記載が下記の点で、特許法第36条第6項第2号に規定する要件を満たしていない。 |
| Translation result without numbering | This patent application does not fulfill the requirement of the provision of Patent Law Section 36(6)(ii) in terms of the item(s) listed below regarding the statement of the scope of claims. |
| Translation result with "理由1_" (Reason1_) | Reason 1 -- this application is not satisfying the requirements which a description of Claims is the following point and are specified to Patent Law Section 36(6)(2). |

Usage of variable for fixed expressions

Standardizing numbers and fixed expressions in the documents by using variable.

examples

Claim 1, Claim 2 → Claim "n"
July 7, 2009 → "n"

Analysis of free description part

Finding frequent patterns by analyzing last few words, statistically. (extending word by word)

examples

不明瞭である。
が 不明瞭である。
技術的範囲 が 不明瞭である。
その 技術的範囲 が 不明瞭である。

← Analyzing by extending word by word

Extraction of phrases and technical terms

1,000 frequently-used phrases and technical terms are examined.



Approaches for Improved Translation Result

Replacement of MT engine

Technical term & IP dictionaries

| | <u>Before Mar. 2009</u> | | <u>After Apr. 2009</u> |
|-----------------------------|-------------------------|---|------------------------|
| Technical Term dictionaries | 1.83 million words | → | 2.50 million words |
| Total number of IP terms | 100,000 words | | 200,000 words |

Translation memories

Before Mar. 2009

Industrial Property terms: 6,000

After Apr. 2009

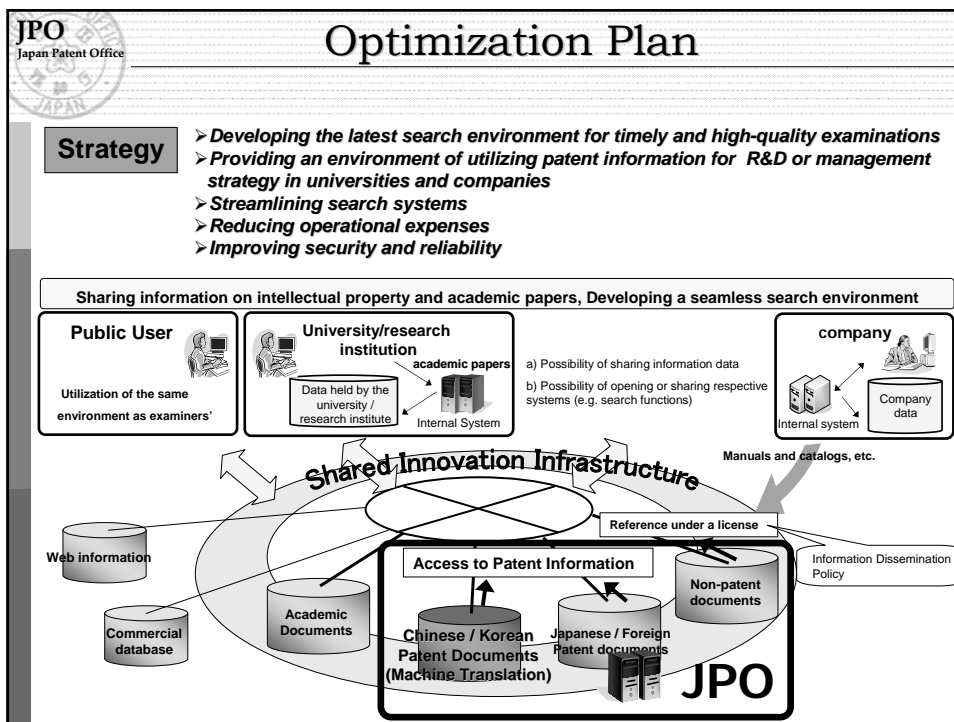
Industrial Property terms: 20,000

Basic terms / Business terms: 150,000

Technical Terms: 400,000

(Chemical / Medical, Biotech. / Machine, automobile / electronics)

Excerpt from brochure of the product



JPO
Japan Patent Office

Optimization Plan

Multi-Lingual Machine Translation

Research on Cross-language information retrieval

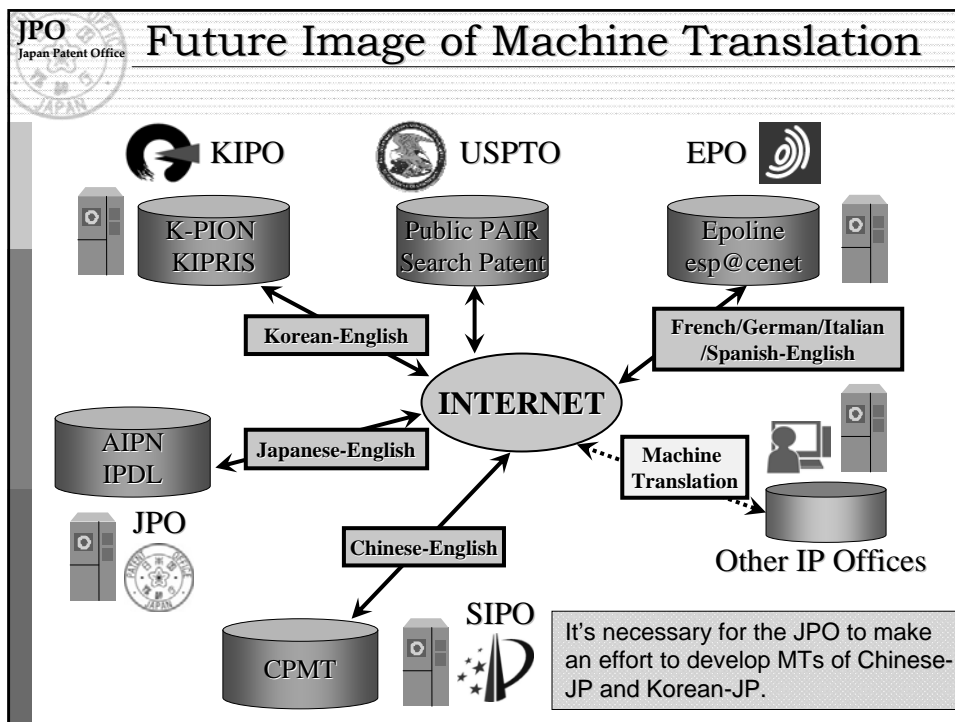
The JPO conducted research on the development of the next-generation search system regarding cross-language information retrieval system.

- Evaluation of the effectiveness of the two translation methods (Keyword translation, Contents translation)
- Evaluation of translation results (Korean-Japanese, Chinese-Japanese)

Korean-Japanese
→ Translated sentences are fairly understandable. (practicable level)

Chinese-Japanese
→ Further development is needed. (especially, enrichment of the dictionaries)

| Types of Dictionary | Chinese | English | Remarks |
|--------------------------------|--------------------|--------------------|---|
| Basic Vocabulary | 250 thousand words | 1.03 million words | Chinese are 1/4 th of English words |
| Technical Term | 238 thousand words | 2.72 million words | Chinese are 1/10 th of English words |
| Automatically registered words | 1300 words | — | — |



JPO
Japan Patent Office

IP5 Foundation Projects

Mutual Machine Translation (lead: KIPO)

Aim: To help the offices overcome the language barrier of patent information and allow greater access to each other's patent information.

IP5 Foundation Projects

- Common Documentation Database (lead: EPO)
- Common Approach for a Hybrid Classification (lead: EPO)
- Common Application Format (lead: JPO)
- Common Access to Search and Examination Results (lead: JPO)
- Common Training Policy (lead: KIPO)
- Mutual Machine Translation (lead: KIPO)
- Common Rules for Examination Practice and Quality Control (lead: SIPO)
- Common Statistical Parameter System for Examination (lead: SIPO)
- Common Approach to Sharing and Documenting Search Strategies (lead: USPTO)
- Common Search and Examination Support Tools (lead: USPTO)