



# Current Status of MT application in JPO

August 30, 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



# History of Patent Information Web Service

- Apr.1996 **JPO Official Website** launched
- Apr.1997 “Patent Abstracts of Japan (**PAJ**)” retrieval service released
- Mar.1999 “Industrial Property Digital Library (**IPDL**)” started on the website
- Mar.2000 “**Full-text Translation**” service released
- Feb.2003 The Official Website renewed
- Oct.2004 IPDL transferred to **National Center for Industrial Property Information and Training (INPIT) Website**
- Oct.2004 “Advanced Industrial Property Network (**AIPN**)” released to IP Offices



# IP Information Dissemination

---

**JPO** : Policy Making on IP Information Dissemination

**INPIT** : Provision of Basic Primary Information

## Databases with Machine Translation

### [IPDL \(Industrial Property Digital Library\)](#)

Offering 71.0 million IP information with search function in Japanese and MT English to public users.

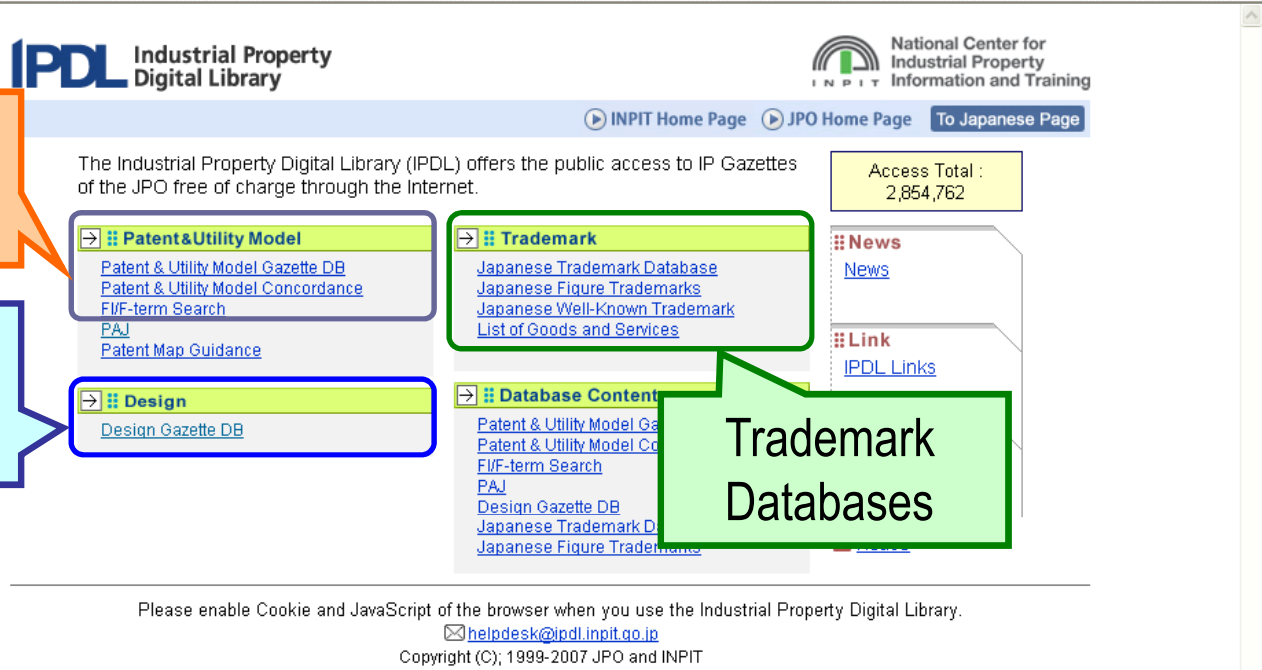
### [AIPN \(Advanced Industrial Property Network\)](#)

Offering patent application information and dossiers in MT English to foreign IP offices.

Patent Databases  
with MT service

Design Database  
with MT service

Trademark  
Databases



The screenshot shows the IPDL website with the following elements:

- IPDL Industrial Property Digital Library** logo and header.
- Navigation links: [INPIT Home Page](#), [JPO Home Page](#), [To Japanese Page](#)
- Text: "The Industrial Property Digital Library (IPDL) offers the public access to IP Gazettes of the JPO free of charge through the Internet."
- Access Total:** 2,854,762
- Patent & Utility Model** section:
  - Patent & Utility Model Gazette DB
  - Patent & Utility Model Concordance
  - F/I/F-term Search
  - PAJ
  - Patent Map Guidance
- Design** section:
  - Design Gazette DB
- Trademark** section:
  - Japanese Trademark Database
  - Japanese Figure Trademarks
  - Japanese Well-Known Trademark
  - List of Goods and Services
- Database Content** section:
  - Patent & Utility Model Gazette DB
  - Patent & Utility Model Concordance
  - F/I/F-term Search
  - PAJ
  - Design Gazette DB
  - Japanese Trademark Database
  - Japanese Figure Trademarks
- News** section:
  - News
- Link** section:
  - IPDL Links
- Footer: "Please enable Cookie and JavaScript of the browser when you use the Industrial Property Digital Library." and "Copyright (C); 1999-2007 JPO and INPIT"

## 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

### Full-text Patent Publication

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

JAPANESE [JP,2006-101591,A]

Mode Index

---

CLAIMS [DETAILED DESCRIPTION](#) [TECHNICAL FIELD](#) [PRIOR ART](#) [EFFECT OF THE INVENTION](#) [TECHNICAL PROBLEM](#) [MEANS](#) [EXAMPLE DESCRIPTION OF DRAWINGS](#) [DRAWINGS](#)

---

\* NOTICES \*

**JPO and INPIT are not responsible for any damages caused by the use of this translation.**

- This document has been translated by computer. So the translation may not reflect the original precisely.
- \*\*\*\* shows the word which can not be translated.
- In the drawings, any words are not translated.

---

CLAIMS

Full-text by MT

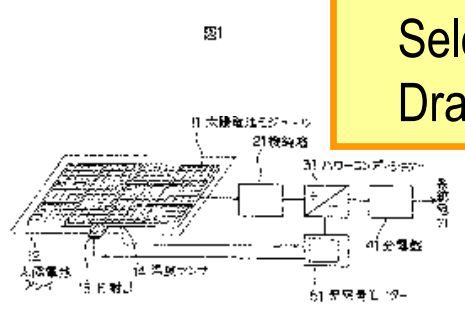
---

[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

Drawing selection

Representative drawing



Selected Drawing

[Translation done.]

## **AIPN**

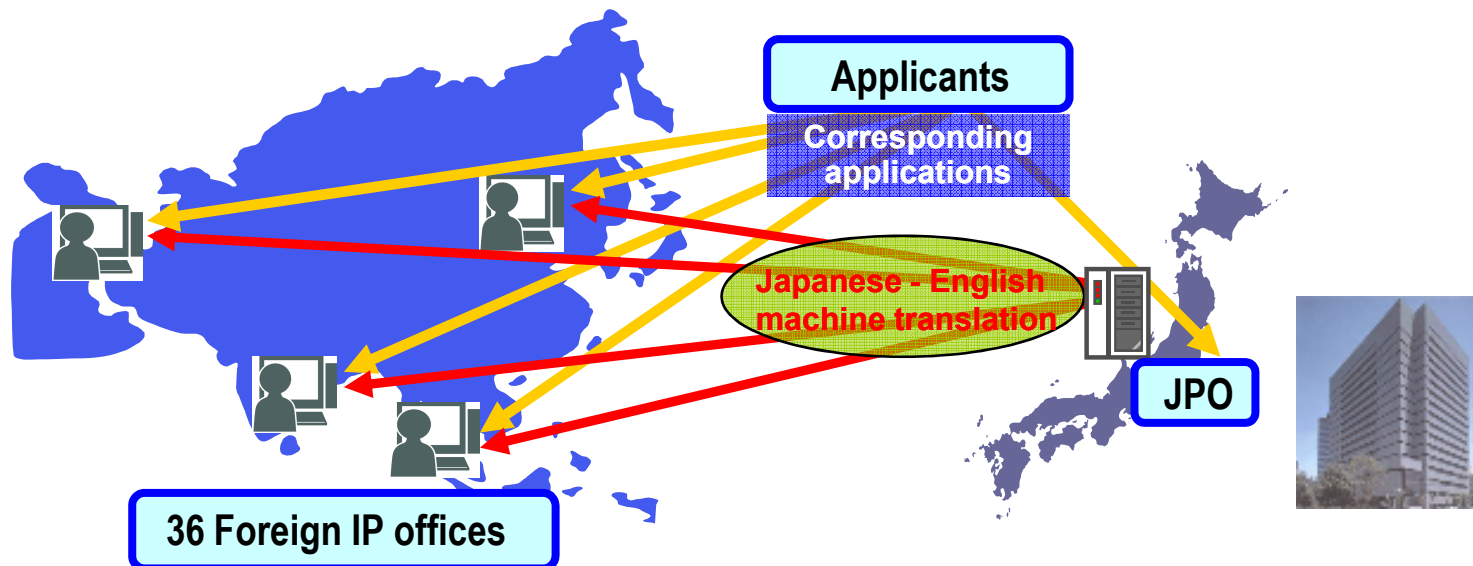
### **(Advanced Industrial Property Network)**

- 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**



### **Advantageous effect**

1. Reducing burden on examiners in foreign IP offices
2. Accelerating quick examination and acquisition of the right



## File Wrapper Document

- 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.]



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 INPIT, and those who drafted this document in the original language are not responsible for the result of the translation.

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

Translated: 16:32:55 JST 06/05/2008  
Dictionary: Last updated 05/30/2008 / 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 (d)0.1-2.0micrometer 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.

Description Document

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 INPIT, and those who drafted this document in the original language are not responsible for the result of the translation.

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

05/2008  
30/2008 / Priority:

Notification of Reason for Refusal

Notification of Reasons for Refusal

application for patent 2001-123456  
No. 15(2003) August 12  
Applicant: AMI, Nobuhiro 9341 3S00  
Attorney: NISHIKAWA, Shigeaki  
Patent Law Section 29(2)

The application should be refused for the reason mentioned below. If the applicant has any argument or other argument should be submitted within 60 days from the date on which this notification is issued.

Reason

The invention(s) in the each claim listed below of this patent application should not be granted a patent under the provision of Patent Law Section 29 (2) for the reason that the claimed invention(s) is/are made by persons who have common knowledge in the technical field to which the invention(s) pertains.

### 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<BR>`

#### 2. Deleting unnecessary spaces

「引 用 文 献 等 一 覧」 → 「引用文献等一覽」  
“The list of cited documents”

#### 3. Pre-formatting

Document number: “yyyyxxxxxx” → “yyyy-xxxxxx”

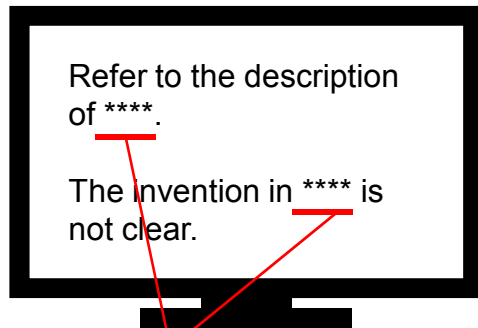
Avoiding Consequent quotation: “bcd)(xyz” → “bcd)(xyz “

# Approaches for Improved Translation Result

## Collection & Registration of unknown words

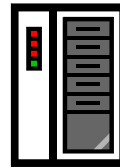
### Collection of unknown words

Translation result (IPDL/AIPN)

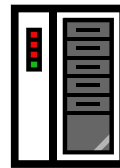


Unknown word is displayed as "\*\*\*\*".

IPDL server

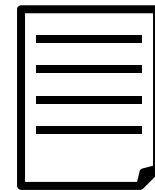


AIPN server



Collection of mistranslation logs

Unknown Words log



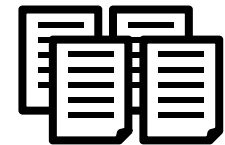
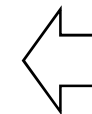
### Adding the meaning, part of speech etc. to the unknown words based on frequency.

Example:

ニッケロセニルーイソープロピル基 → nickeloceryl-iso-propyl  
重合体カテキン → polymer catechin



Dictionary data  
(5,000 words/year)



Unknown words logs

### Registration of the unknown words

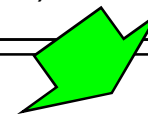
(5,000 words/year)

Provision of dictionary data at marginal cost

(65,000 words : as of March 2009)



Registration to the dictionary



Provision to public



# Approaches for Improved Translation Result

## Mistranslation feedback

JPO has received mistranslation feedbacks from AIPN users.  
(EPO and USPTO)

### Feedback process (once a month)

1. Verification of repeatability
2. Checking up original sentences and adding appropriate meanings
3. Review of added meanings  
by referring PAJ, IPDL and some technical term dictionaries
4. Assessment of side-effect caused by additional words registration
5. Authorization by JPO and INPIT
6. Registration in user dictionary  
(It takes 2 months to complete the process)



# 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

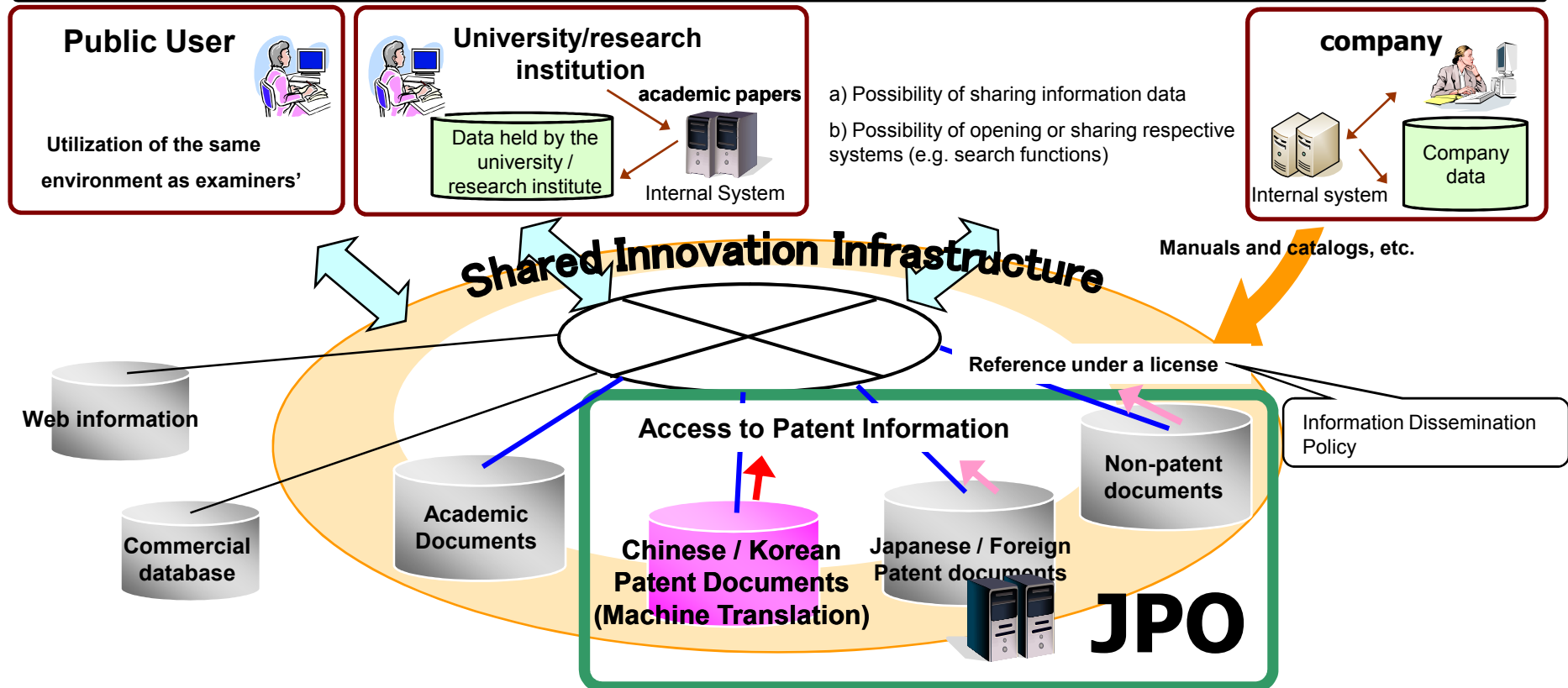
(Chemical / Medical, Biotech. / Machine, automobile / electronics)

# Optimization Plan

## Strategy

- *Developing the latest search environment for timely and high-quality examinations*
- *Providing an environment of utilizing patent information for R&D or management strategy in universities and companies*
- *Streamlining search systems*
- *Reducing operational expenses*
- *Improving security and reliability*

Sharing information on intellectual property and academic papers, Developing a seamless search environment



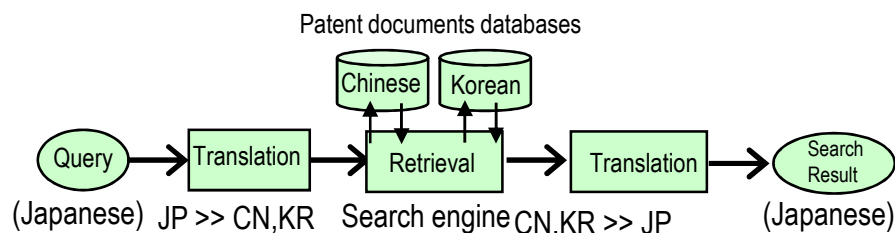
### Research for “Cross-lingual retrieval system”

JPO has a plan to establish “Cross-lingual retrieval system” which enable to search Chinese and Korean documents.

- Evaluation of the effectiveness of the two translation methods (Keyword translation, Contents translation)

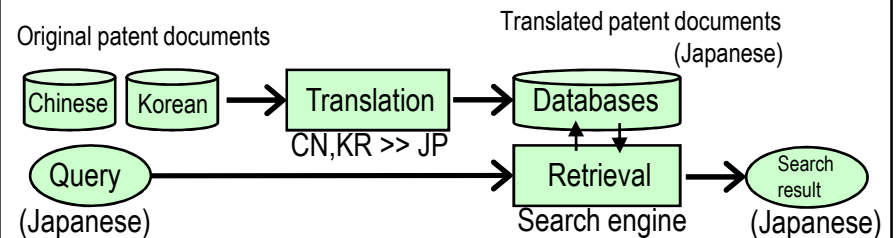
#### Keyword translation

Query in Japanese is translated into Chinese or Korean in each transaction for search.  
Patent documents, written in the original language, are searched by using translated query.



#### Contents translation

Original patent documents are translated into Japanese and stored into database in advance.  
The translated databases are searched by using query in Japanese



- Evaluation of translation results (Korean-Japanese, Chinese-Japanese)

#### Korean-Japanese

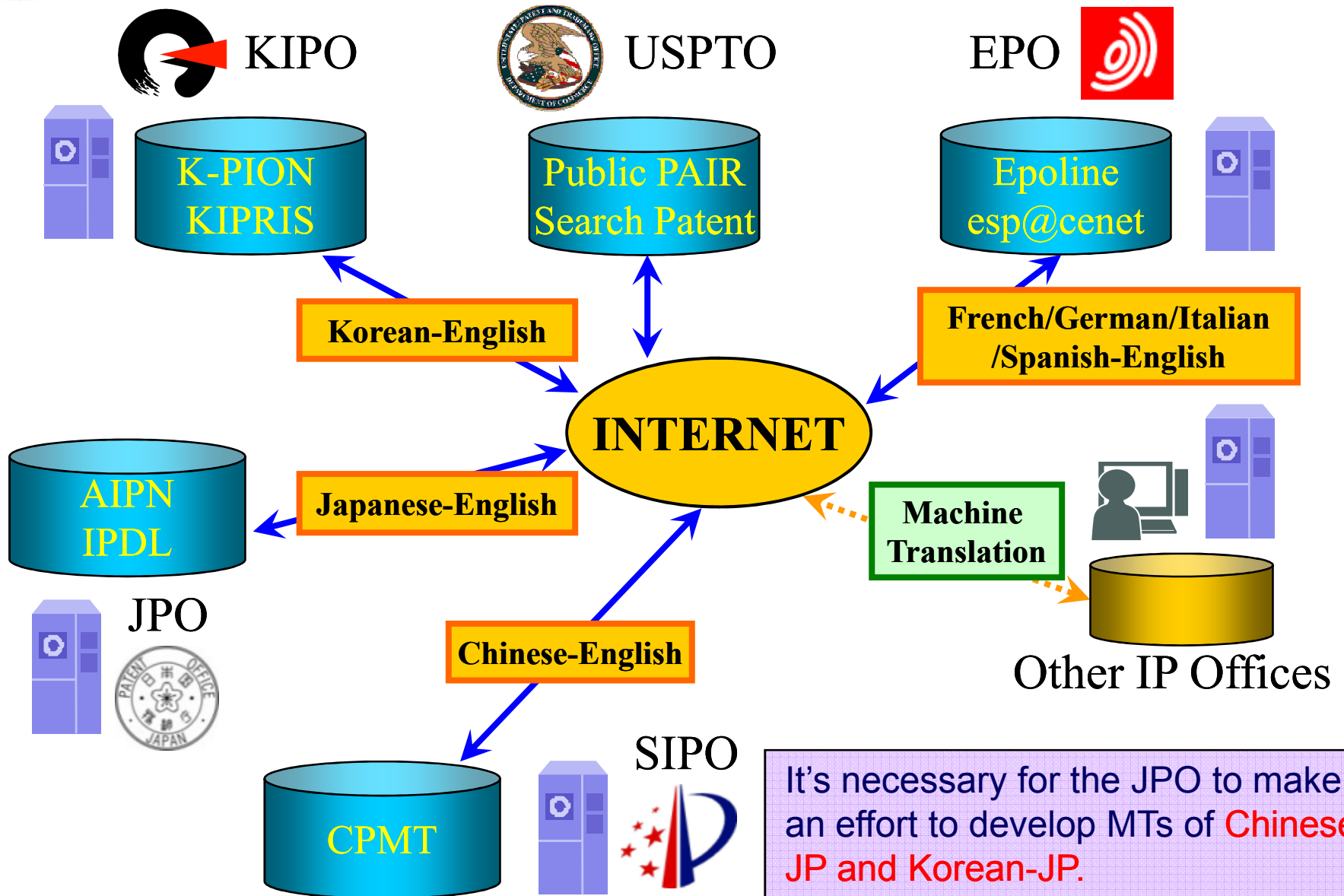
>> Translated sentences are fairly understandable.  
(practicable level)

#### Chinese-Japanese

>> Further development is needed.  
(especially, enrichment of the dictionaries)



# Future Image of Machine Translation



It's necessary for the JPO to make an effort to develop MTs of **Chinese-JP** and **Korean-JP**.

## 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)



*Thank you*