



# Challenges of the Asian IP information for non native speakers



Sandra Bailly

Information Retrieval Facility Symposium - 2008

# 2008-2009 P&G Billion-Dollar Brands



# Background

- More than 50% of the patent applications published WW are written in Japanese, Korean or Chinese
- Japan is the world's second biggest PCT user
- Domestic Korean applications doubled from 1997 to 2006
- Domestic Taiwan applications increased fivefold from 1997 to 2006





Is automatic machine translation the  
answer?



# Translation Systems available to the IP community:

- Human Translation
- Machine translation
- Human assisted translation
- Cross-lingual machine translation



# What is available to use in the IP information area?

Asian data coverage and automatic translation tools are consistently improving from both:

- Patent offices
- Commercial providers



# What is available to use from the Patent Offices?

## Korean Patent Office: KIPRIS

- English abstract from 1979
- New since 2007: full text cross lingual patent and utility model retrieval tool (Payable)

## Japanese Patent Office: IPDL

- English abstracts for patent applications from 1976 on
- Machine translated full text from 1993 on



# What is available to use from the Patent Offices?

Patent search at the Chinese Patent Office (SIPO), China Patent Database (CNPAT), China Intellectual Property Net -CNIPR:

- English abstracts for patents (human) and utility models (MT) from 1985 on
- TEST: Chinese – English machine translation



# What is available from the commercial providers?

## Translation of Asian IP information:

- Human
- Machine translation
- Human assisted translation

## Structured data:

- Index Terms
- Codes
- Classifications
- Assignee names

=> Value added information



# An example

Will I retrieve this patent from  
AKZO?

Title: [MT] Method and apparatus of the degree of roughness used for measuring and is filmed

Application Number: 200580042756

Application Date: 2005.12.13 Publication Number: 101076833  
Publication Date: 2007.11.21

IPC: G06T7/40(2006.01)I;G01J3/51(2006.01)I;G01J3/46(2006.01)I

Applicant: [MT] **Akezuo Nobel Coatings International Inc**

Inventor: [MT] Njo Swie Lan Priority Information:  
US2005065447820050222;EP2004007838320041214

Abstract: [MT] The patent refers to the field of 'image data processing or generation'. A method used for using for analyzing and comprising the **visual degree of roughness** that **films of the pigment** of effect through the measuring set, this measuring set includes reflecting the inboard wall and open end lumen of sample, this device also includes: The illuminator, are used for lighting the stated lumen; etc....

<http://english.cnipr.com/enpat/search/tableSearch.do?method=tblSearch>

Paint film?



# Another example

The screenshot shows a Windows desktop with Adobe Reader and Microsoft Internet Explorer open. The Internet Explorer window displays a search result page with the following text:

Address: <http://www4.ipdl.inpit.go.jp/Tokujitu/PAJdetail.ipdl?N0000=60&N0120=01&N2001=2&N3001=1>

JAPANESE [JP,2002-020783,A]

either [at least / whose ] residue R<sup>6</sup> or R<sup>7</sup> is H.  
[Claim 13] The fragrance outsider compound according to any one of claims 1 to 12 residue R<sup>6</sup> and whose R<sup>7</sup> are H.  
[Claim 14] Residue R<sup>6</sup> and R<sup>7</sup> are H, and R<sup>1</sup> - R<sup>5</sup> independently, H, -NO<sub>2</sub>, straight-chain-shape, or branched chain-like C<sub>1</sub> - C<sub>6</sub>-alkyl. The fragrance outsider compound showing C<sub>1</sub> - C<sub>6</sub>-alkenyl, C<sub>1</sub> - C<sub>6</sub>-alkynyl or C<sub>1</sub> - C<sub>4</sub>-ARUKOKISHI according to any one of claims 1 to 13.  
[Claim 15] Fragrance ketone expressed with the formula II 1-phenyl-ethanone, 2,4-dimethylphenyl ethanone, 1-[4-(1,1-dimethylethyl)-2,6-dimethylphenyl]-ethanone, 1

The PDF document in the background shows the following numbered list items:

13. Fragrance precursors of formula I according to one of the claims 1 to 12, wherein the residues R<sup>6</sup> and R<sup>7</sup> are H.
14. Fragrance precursors of formula I according to one of the claims 1 to 13, wherein the residues R<sup>6</sup> and R<sup>7</sup> are H and R<sup>1</sup> to R<sup>5</sup> represent independently H, -NO<sub>2</sub>, linear or branched C<sub>1</sub>-C<sub>6</sub>-alkyl, C<sub>1</sub>-C<sub>6</sub>-alkenyl, C<sub>1</sub>-C<sub>6</sub>-alkynyl or C<sub>1</sub>-C<sub>4</sub> alkoxy.
15. Fragrance precursors of formula I according to one of the claims 1 to 14, wherein the fragrant ketone of formula II is selected from 1-phenyl-ethanone, 2,4-dimethylphenyl-ethanone, 1-[4-(1,1-dimethylethyl)-2,6-dimethylphenyl]-ethanone, 1-(4-tert-butyl-3,5-dinitro-2,6-dimethyl)-ethanone and 1-(4-methoxy-phenyl)-ethanone.



- Lower the barriers to access Asian IP information.
- Great tools to facilitate my understanding – give an overview of the technology – help to pre-select the information
- Not fully reliable on the information searched and so retrieved



# Language translation software challenges

- Sentence structure
- Context of the sentence
- Concepts
- Grammar
- Spelling
- New words
- Expressions
- Chemical name
- OCR
- ...



# IP Researcher challenges

- Be confident in the query : expressions, words, spelling, chemical names
- Get consistency and trusted data, eg assignee names
- Get confidence in the system



# Maybe next?

One multi-lingual tool with queries and results in the native language of the IP researcher



謝謝您的注意

谢谢

Thanks for your attention

Touching lives, improving life. *P&G*

ご清聴ありがとうございました

