INTRODUCTION

• Overview of Derwent World Patents Index (DWPI)
• Processing Patent Assignees
• Patent Assignee Codes
• Ongoing Projects
OVERVIEW OF DWPI VALUE ADD
DERWENT WORLD PATENTS INDEX® (DWPI)

• 52 data sources

• Deep coverage:
  • Pharmaceuticals from 1963
  • Agrochemicals from 1965
  • Polymers from 1966
  • All chemistry from 1970
  • All technologies from 1974

Contains >30 million unique inventions containing details of 64 million individual patent documents

Who is using DWPI?  IP Attorneys, IP Researchers, R&D, Information Professionals, Portfolio Managers, Competitive Intelligence and Patent Examiners (in 40 offices worldwide)
Global support in the creation of the Derwent World Patent Index value add
DERWENT WORLD PATENTS INDEX® PROCESS OVERVIEW

Multiple Data Sources | Multiple Languages

- Full Text, Gazettes, 3rd Party Translations
- Rules Engine
- Aggregation & Normalization
- Editorial Correction
- Machine-Assisted Translation
- Abstracting & Classification

Also delivered via FLD channels

Pegasus

HUMAN INTERVENTION
DWPI STATISTICS

• Process 87,000 patents per week
  – 50% are new inventions
  – Data corrections/review required for around 20% of patents

• 900 Editorial staff

• ~2.5m new records created per year
DWPI ASSIGNEE DATA
WHERE DOES ASSIGNEE DATA FIT IN TO THE EDITORIAL PROCESS?

Preliminary processing
Bibliographic Data
Patent Family
Basic/Equivalent Search
Patent Assignee Codes
Non-Convention Equivalent
Sectioning
Classification

Quality Control – Data Integrity

Product Creation (DWPI)
DWPI VALUE ADD - STANDARDISED PATENT ASSIGNEE NAMES

1. **In-house Company Dictionary**
   - Company names are checked against an internal database
   - Contains more than 1.8 million entries

2. **Patent Assignee Name Standardized**
   - Error corrections (spelling, transliterations)
   - Abbreviation and Standardization

3. **Patent Assignee Codes (PACO)**
   - ID of patents owned by one and the same legal entity
   - Name variations, subsidiaries, joint ventures, mergers, depts.
   - Consistently applied and maintained through an editorial process
HOW DOES DWPI HELP?

COMPANY NAMES

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Jahr et al.</td>
<td></td>
<td>Date of Patent:</td>
<td>Jun. 8, 1999</td>
</tr>
</tbody>
</table>

- **In DWPI, 21k+ company names are standardized, including subsidiaries**
- **Primary Examiner**—Ayaz R. Sheikh
- **Assistant Examiner**—Ario Etienne

**Original patent data corrected in DWPI**

**Ensure that searching & monitoring by company returns all results**

Corrected to **Hewlett Packard Co** and given standard DWPI Patentee Code: **HEWP**
Patents are not a faultless source of data.

Some errors may be easy to spot such as “Fabebook”, but many are not so obvious.

A searcher would miss this patent if searching by patentee name or patentee code (PACO).
One variation of a system for buttering and toasting bread buns includes: a butter stage including: a vessel containing liquefied dairy fat and defining a permeable surface, a pressure supply fluidly coupled to the vessel, and a first paddle assembly configured to catch a bun crown dispensed into the butter stage, to compress a sliced face of the bun crown
PATENT ASSIGNEE CODE (PACO)

- Each assignee in DWPI also has a Patent Assignee Code. We have four separate series of PACO
  - Standard codes, these end -C. e.g. SIEI-C
  - Non-standard codes, these end -N e.g. ACME-N
  - Individual’s name, these end -I e.g. SMIT-I
  - Russian companies, these end -R e.g. UNVO-R
PROCESSING ASSIGNEE DATA

• Company names are checked against an internal database (currently over 1.8 million entries)

• Patent Assignee Codes added for recognised companies

• For new company names (not previously recognised), a new code is assigned and added to the internal database
WHY DO WE USE PACO

- PACO helps DWPI users to:
- Distinguish between different companies with similar names

Eg

Merck KGaA
PACO: (MERE-C)

Merck & Co., Inc.
PACO: (MERI-C)
WHY DO WE USE PACO

- PACO help customers to:
- Identify subsidiary companies which do not carry the parent company name
- Eg

- Intel Corp
- PACO: ITLC-C

- McAfee Inc
- PACO: ITLC-C
ASSIGNMENT OF STANDARD PACO

• When a patentee has 500 patents in DWPI then we assign a standard PACO
  - Examples include Siemens AG - SIEI-C, Pfizer Corp - PFIZ-C, Robert Bosch GmbH - BOSC-C, Glaxo Inc - GLAX-C

• Identified by interrogation of internal database

• Patentees researched using internet searches and company websites

• Use research to identify subsidiary companies
  - Honeywell Inc, Honeywell Corp, Honeywell GMBH, Control Systems Ltd, Pittway Corp all have the code HONE-C
NAME CHANGES, MERGERS AND ACQUISITIONS

• When a company changes its name then the old standard PACO will continue to be used

• If two companies with standard codes merge, the old PACO will be used if records are filed using the old names

• DWPI does not automatically assign a new standard PACO when two large companies merge etc. Each case is assessed individually

• PACO are not generally changed retrospectively as the assignment of patent rights are not tracked by DWPI
CHALLENGES IN IDENTIFYING CORRECT ASSIGNEE & PACO

- We receive over 87,000 patents each week
- 11,000 of these are reviewed by analysts to check patentees & PACO
  - Where the dictionary does not recognise a company, it is checked by the analyst
  - If it can be identified as a member of a known company, the analyst will assign the correct PACO
  - All new patentees receive a non-standard PACO
CHINESE PATENTEES

• Large volumes of Chinese patentee names create a challenge for DWPI

• We rely on transliterated name

• Construction of patentee name is usually preceded by a city or province

• Chinese universities have been identified as starting with City/Region name followed by speciality

• Naming convention of Chinese universities differ from western universities
CHINESE ASSIGNEE PROJECT

• To create new “standard PACO” for Chinese companies

• Identify & research Chinese companies with more than 500 records in DWPI

• Create new PACO for these

• Correct all backfile DWPI records for the new PACO

• CURRENT PROGRESS - 413 New PACO generated
  – 772,000 records corrected & being released to product
PATENT ASSIGNEE CODE LOOK-UP

- Look-up tool on the Thomson Reuters website
- Browse tool within Thomson Innovation
- PACO Thesaurus on STN
CONCLUSION

• Processing assignee data is a significant challenge
  – High volumes of data, increasing annually
  – Non-English language data
  – Grouping into Patent Assignee Codes

• So what’s next?
QUESTIONS?
Get insights you would otherwise miss and make more confident decisions by searching the authoritative and accurate patent data in DWPI.

Global patent coverage you won’t find anywhere else — In English
Patent data made accurate by industry experts, helping you find what’s important to you
 Reliable and enhanced data, featuring clear titles, abstracts summarizing the novelty of the invention
Precision indexing provides unsurpassed searchability