

# AI IN PATENT RESEARCH

Ed White, Director Patent Analytics

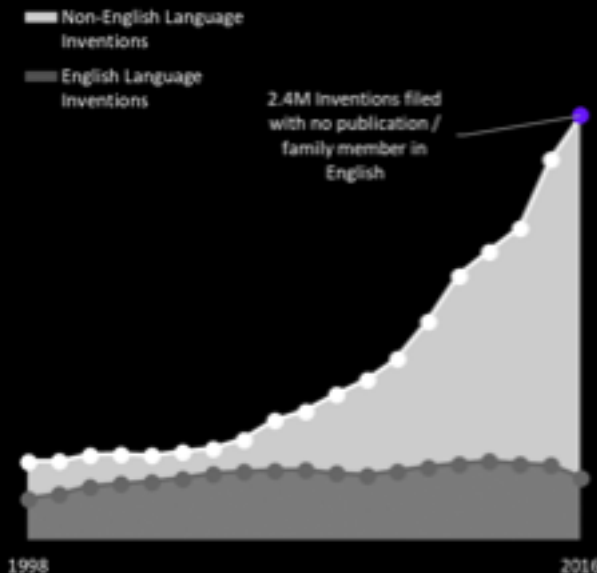
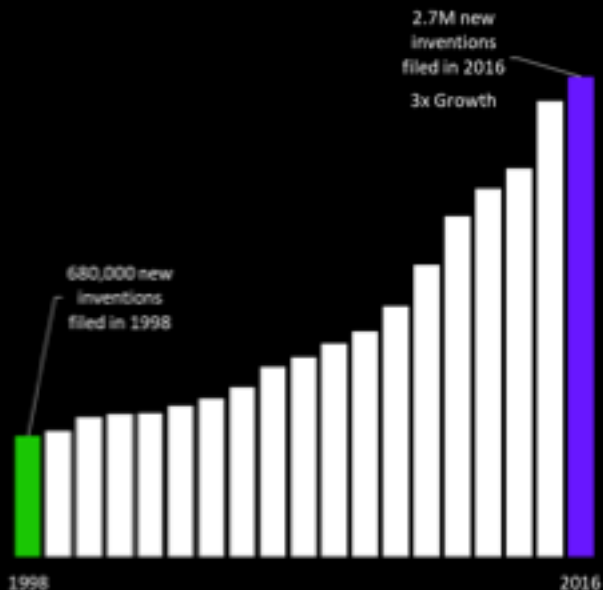
**Derwent**  
*Powering IP Innovation*

 **Clarivate**  
Analytics

More data than ever...

...not in English...

...presented in jargon



“What is claimed .... a method comprising... a computer readable medium”

Earliest Priority Year Assignee / Reassignment

G06K9/00791

The present application is a continuation of U.S. patent application Ser. No. ...

Patent Co-operation Treaty

- Information cannot be read or curated by hand – at least by a user
- Even if you wanted to, the raw information is not in a language you can read
- The complexity of patent data limits the level of insight available to the level of expertise of the user
- The promise of powerful global, deep technical intelligence is limited in utility and audience

### **The Goal:**

**Increase the number of professionals that can lean their decision-making on patent data sources**

# Where is automation in use within Derwent?



Construction



Cleaning



Curation



Metrics



Visualization



Insight

## Editorial Production Automation

Duplicate Invention ID and Error Correction Algorithms

Expertise Routing and Basic Tech Classification

Human-trained Machine Translation Engines, specific to Patent Data

## User Aids

SmartSearch

Optimised and Ultimate Assignee (coming soon!)

Strength Scoring  
(coming soon!)  
Predictive Legal Data

ThemeScape

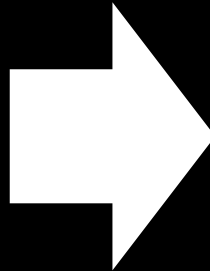
## Gold Standard Data

Hand-crafted, providing training data source for algorithmic extension

# Augmented Expertise – leaning on data



Smart Search



Leaning on 40 million English-language and manually classified Derwent Families

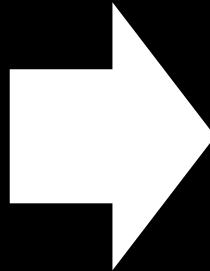
Iteratively expanded semantic algorithms, via DWPI fields and Indexing

Searches fields you don't know about, returning tight datasets beyond your expertise / available time

# Augmented Expertise – leaning on data



Optimised, Probable  
and Ultimate  
Ownership



Takes all ownership data  
available: DocDB, Raw Record,  
DWPI Editorial Assignee Codes  
and more

Leans on known corporate  
hierarchy data, latest  
reassignment data and implied  
reassignments in the family

Predicts who owns the  
invention, even if only an  
inventor listed

# Augmented Expertise – leaning on data



Uses of Graphene, 2017, via Derwent ThemeScape

Standard (and quite evolved)  
natural language algorithms

Applied to hand-crafted,  
contextualised information, e.g.  
Derwent “Use” field

Provides clean information and  
insight on state of the art

# Where are we going next? Where should we go next?



Construction



Cleaning



Curation



Metrics



Visualization



Insight

Distil diverse complex data so it's open to a much wider audience

Assisted Patentability, Validity, FTO, Evidence of Use?

## Gold Standard Data

Hand-crafted, providing training data source for algorithmic extension



# Further ahead – possible game changers, skill set transformation

- Automated Patent Examination?
- Automated Patent Drafting?
- Automated inventing?
  
- Who uses patent information in 10 years?
- How did they get the information? Organise it?
- How does the skill set of the current patent analyst change? More expertise? Less?

# Q&A