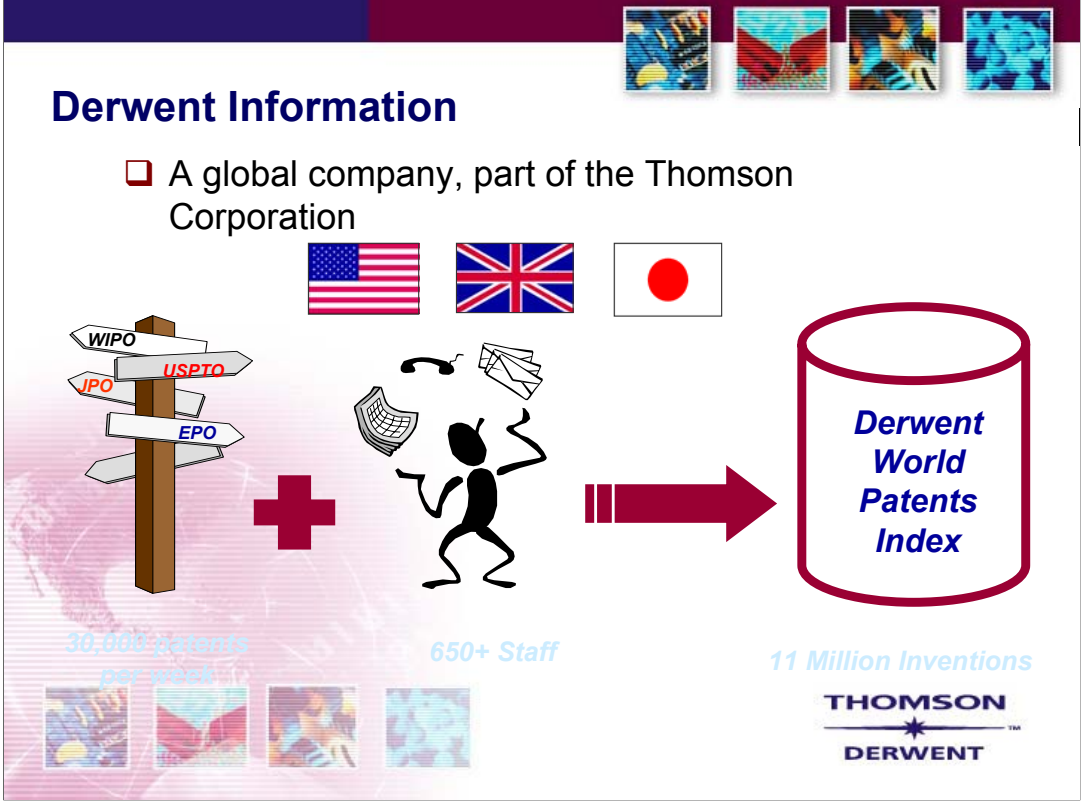
Good afternoon everyone.

First of all, many thanks to OEPM and the Ministry of Science and Technology for their kind invitation to speak today. That appreciation comes from my company, Derwent, and from me personally.

Just as the legal systems for intellectual property protection themselves are complex, so are all the issues surrounding their information and documentation.

Having first discovered patents as a research chemist joining industry from university some 40 years ago, and now spending most of my time with our users in patent offices, I am certainly still learning.

Patent information must be about the users. Those of us involved with running the information systems need to work together for their benefit



We're a global company, part of the International Thomson Organisation, with over 700 employees located worldwide in London, Manchester, Glasgow, Tokyo and Washington DC.

We take patent documents from 40 patent issuing authorities around the world - on average 30,000 documents per week - and extract the business critical information from these complex documents through a process of translation, standardisation, abstracting and indexing.

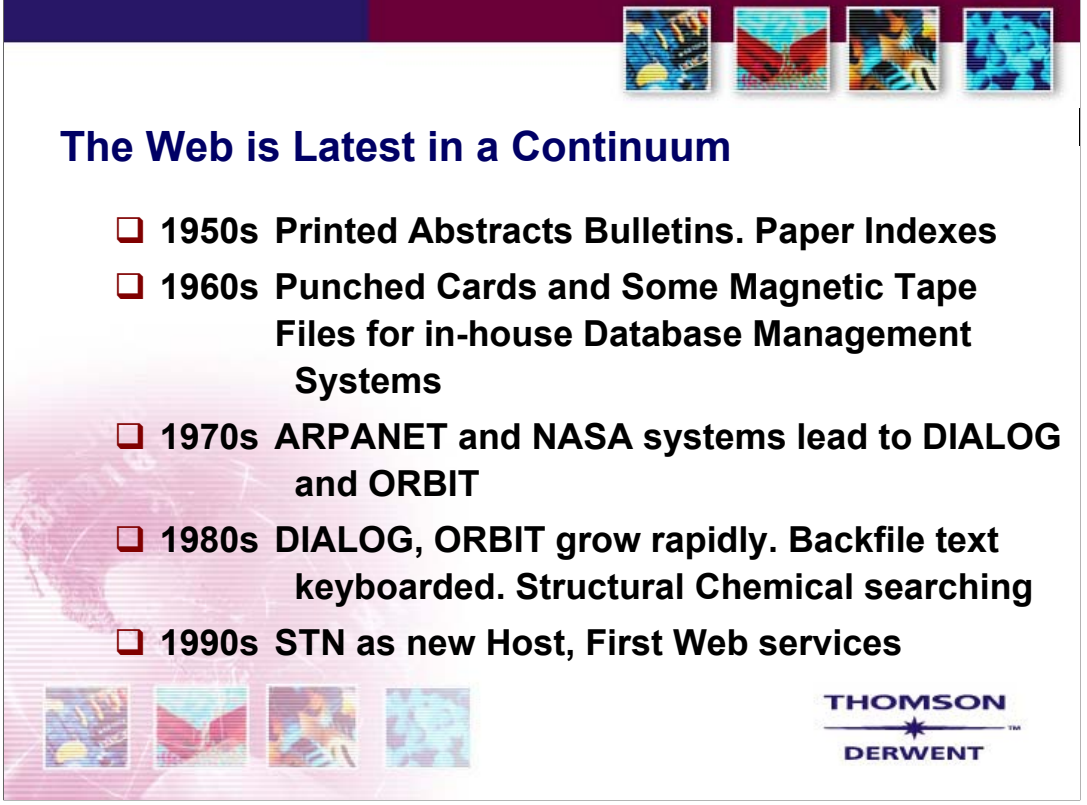
The major output is the Derwent World Patents Index - a database that now contains 11 millions records representing over 20 million patent filings since 1963.

Introducing Derwent

- World's leading supplier of value-added patent and scientific information

- Derwent World Patents Index
  - World's largest value-add database
    - 11 million "basic patents" plus equivalent filings from other issuing authorities
    - Covers 40 issuing authorities including:
      - USPTO
      - JPO
      - Major European Offices

THOMSON DERWENT

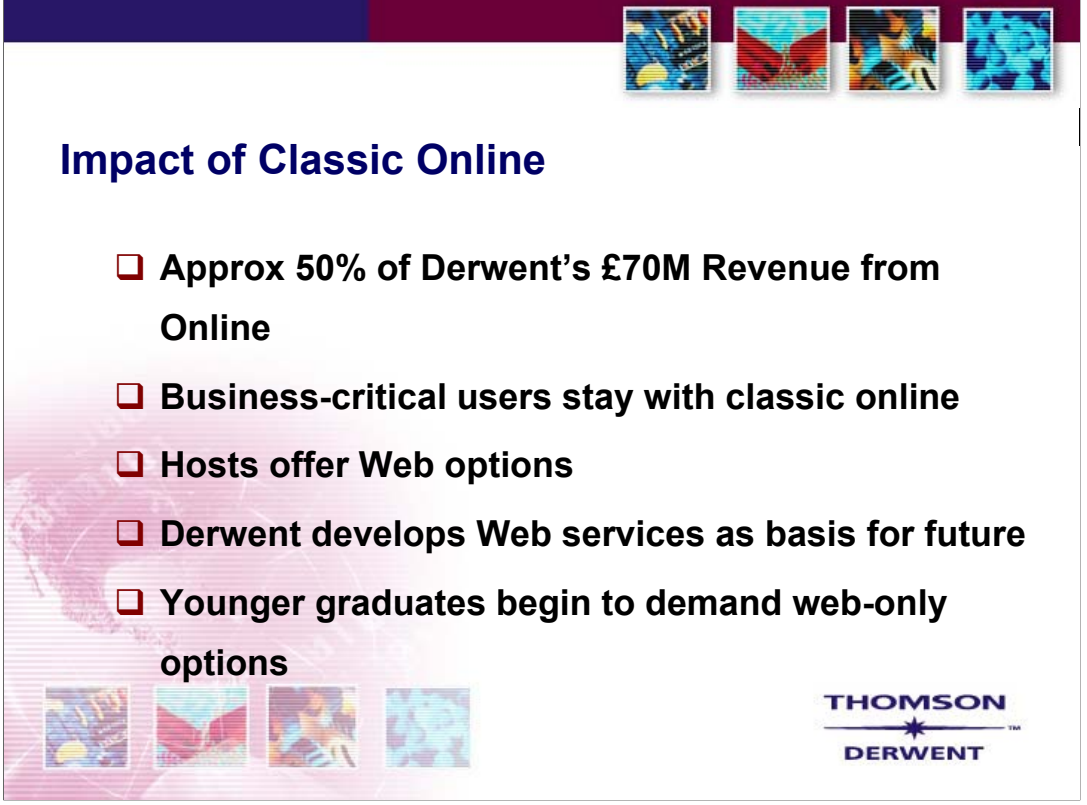


As an information scientist, I really have to view the Web as the latest in a string of developments, each one of which seems to be more absorbing and useful than the last.

However, we are a long way from being able to have instant wisdom.

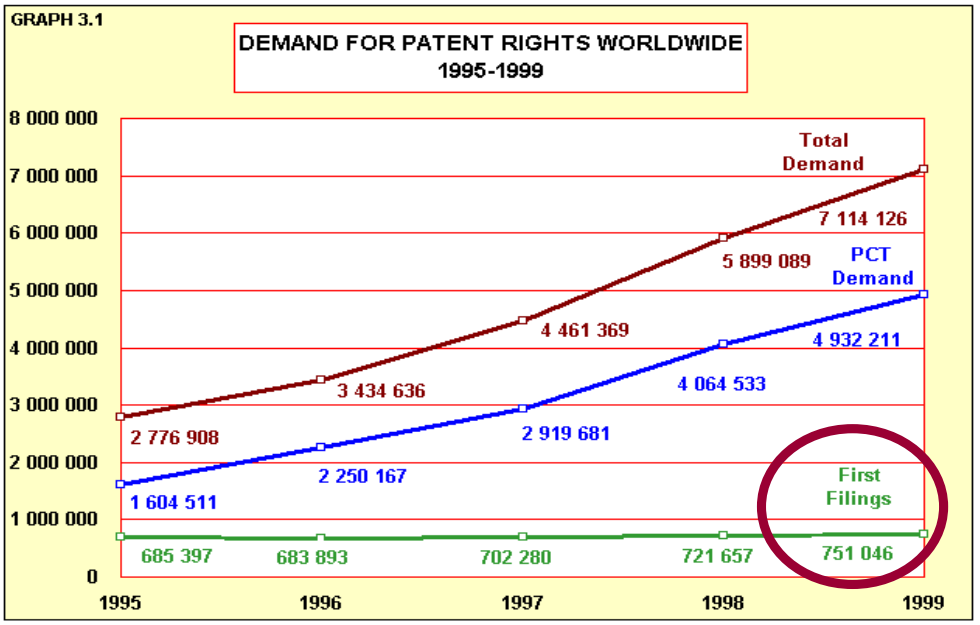
We are equally far from any thought of a single system fitting everyone's needs.

Innovation applies as much to our IP information systems as it does to technology in general



From our Derwent viewpoint, this is what we see happening at this moment.

Without doubt most end-users with occasional use prefer the intuitive web systems.

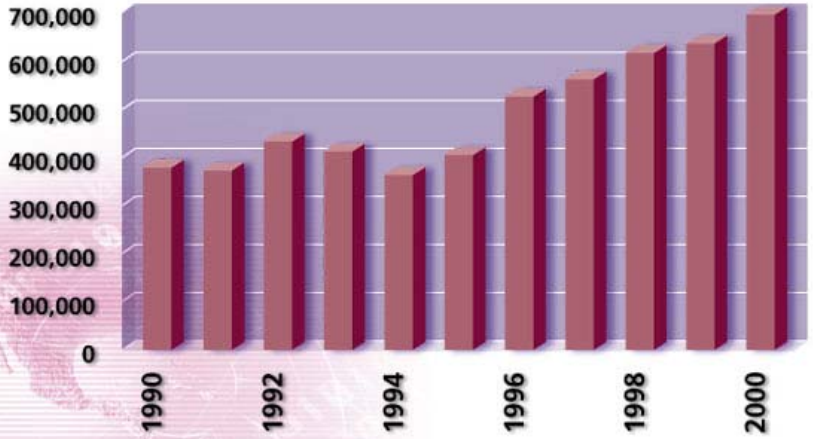


Source = 2000 Trilateral Statistical Report

This graph is from the Trilateral Report, which is on the EPO web site.

Notice the increase in the numbers of patent documents and the relationship to numbers of actual inventions as defined by “first filings”

Annual increase in patented inventions



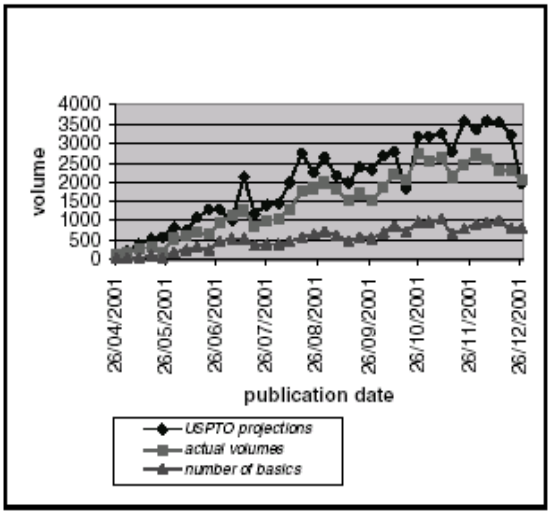
Number of new patents from 40 patent issuing authorities added to Derwent World Patents Index over the past 10 years

THOMSON DERWENT

One of the important things which Derwent does is to cover 40 “countries” and organise the records so that, more or less, you have one record per invention.

This year we expect to cover 1.5 million new inventions

Growth in New US Applications



THOMSON DERWENT

By far the largest impact in 2001 to 2002, and continuing has been from the new US Applications.

It is the cost of covering such volumes of information, making abstracts and indexing,which makes Derwent relatively expensive.

For us, new information policies of patent offices have not changed the economics, but streamlining production processes has helped to reduce the scale of increases.

One of the re-discoveries of the Web age, is that content is important.

Basically, Derwent's services, right from the beginning, have been influenced continually by customer needs

Finding the right patent

- Millions of documents X
- Various languages X
- Deliberately obscure X
- Hundreds of pages X
- Legal jargon X
- Unlikely to be instantly to hand X

THOMSON DERWENT

These are the obstacles that all users are faced with in trying to get the right answers.

Worldwide, over 30-40 million patents have been published to date, and over four million* are currently in force.

The patents may be filed at one of over 40 patent offices, in a variety of languages.

Because patents are usually filed well before the inventions they describe are ready for market, the patent experts who write them frequently try to obscure the novelty of the invention, to reduce the chance of competitors finding the information.

To add to these problems, some patents contain well over 200 pages of technical information, or even thousands of pages in the case of genetic sequences

And they are written using legal terminology unfamiliar to those more used to scientific or technical text.

The sheer numbers of patents published also make it unlikely that an original patent document of particular interest will be "on the shelf" for instant browsing

Adding value to patent information

- ***Abstract*** **key information**
- ***Remove*** **legal jargon**
- ***Add*** **subject indexing**
- ***Assign*** **patentee codes**
- ***Select*** **the key drawing/structure**
- ***Write*** **completely in English**

THOMSON DERWENT

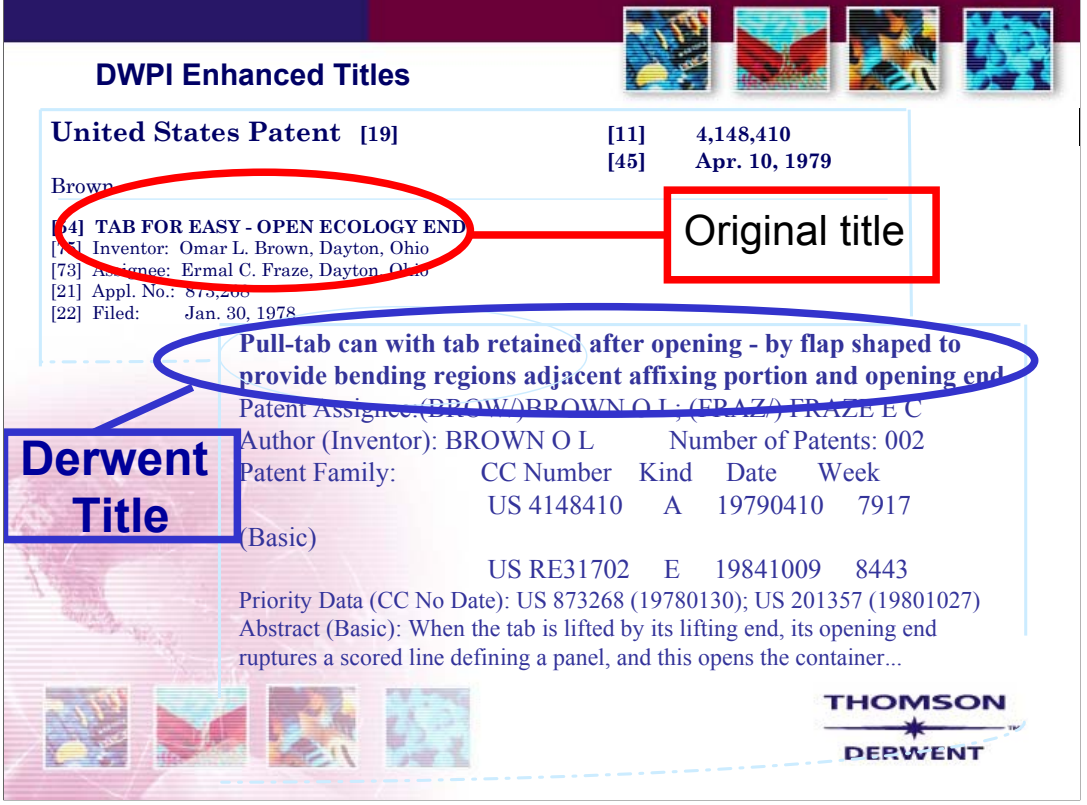
One of the main ways we add value to this information is by extracting the key information from the patent.

Our editors create an enhanced title and **abstract** which summarises the most salient points and key technologies described in each patent. We remove any **legal jargon** so that the technical content of the patent is easily understood, and make it easy to find related patents by grouping them together in "patent families"

We add **subject indexing** to our patent records to enable precise, technology-specific searching of our data - more about this later.

We also assign **patentee codes** to the 21,000 organisations worldwide that patent frequently, enabling rapid searching for applications from specific companies.

Because there is no standard language for patents, we provide all this information **in English**, whatever the language of the original patent document



Both titles and abstracts are specially written by us - and that makes even simple word searching much more effective - not to mention spotting a relevant item from a long list

Derwent Indexing

Bibliographic

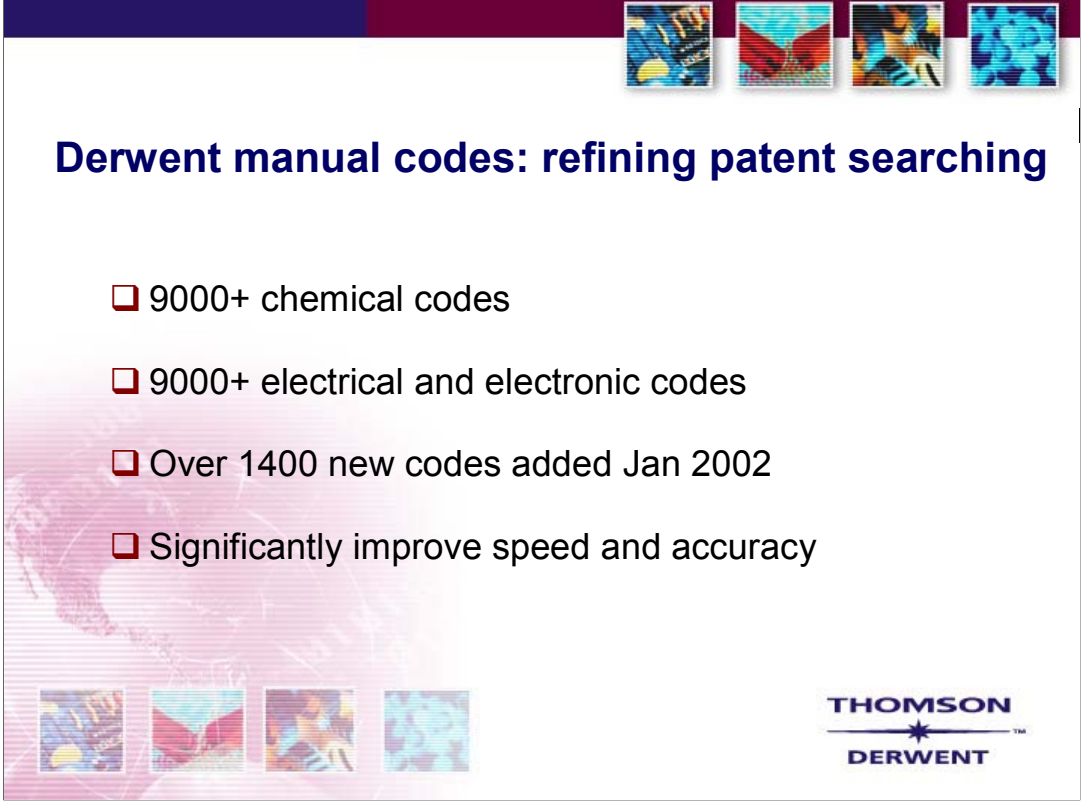
- **Patentee Codes**

Technical Content

- **Linear Hierarchies (like IPC)**
  - Derwent Classes
  - Manual Codes
  - Genetic Sequences
- **Fragmentation Codes**
- **Graphic Structures**
  - Markush DARC (MMS)
  - ISIS/MDL

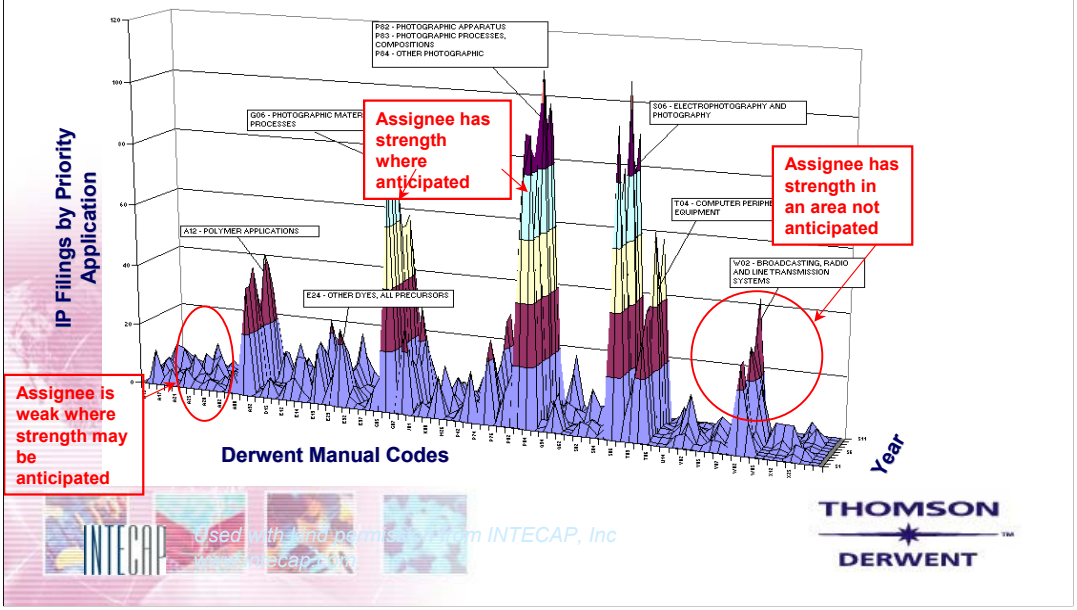
THOMSON DERWENT

This is just a summary of current types of indexing which are also there to make retrieval easier



Easiest to use are so-called Manual Codes

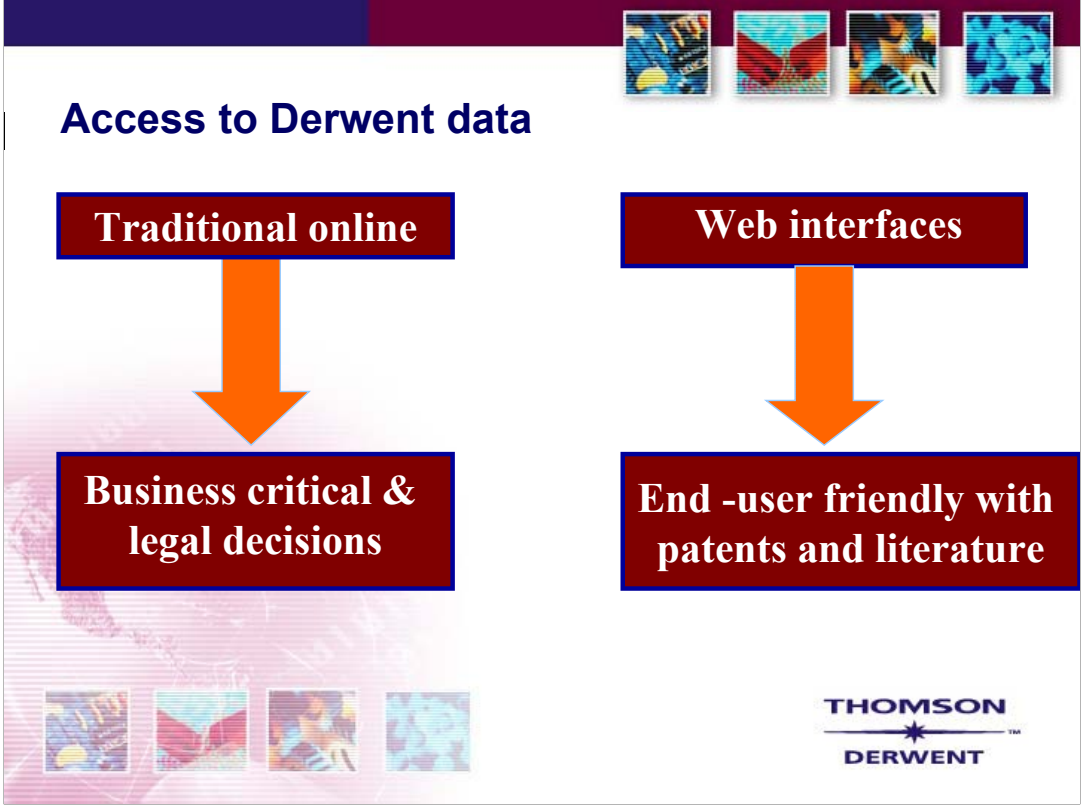
IP Profiling: Assignee Activity over Time using Derwent Classification



Example of how you can use Manual codes for analysis of an organisation's patent portfolio…by courtesy of InteCap who specialise in Intellectual Asset Management studies

"This is a straightforward example of IP profiling and, I think, provides an excellent starting point for analysis. In this case, showing one assignee's filing activity, by Derwent class, over time, presents us with a chart that helps an analyst identify areas where the assignee has strength where we would expect to see strength. We can also see that there are areas that we expected strength and yet the assignee did not have that much activity. And we can quickly determine where the assignee has strength where we did not anticipate such activity.

Finding the surprises are what make a chart like this useful. This also becomes a nice communication vehicle to non-IP professionals since it is visual in nature and can provide information at a high level for quick assessment and understanding for managers with limited time."



Customers like to have a choice, and always will - and each has different needs.

Classic online (Questel, STN and DIALOG) is very important for regular users with demanding searches, and it continues to develop with Web interfaces becoming more popular.

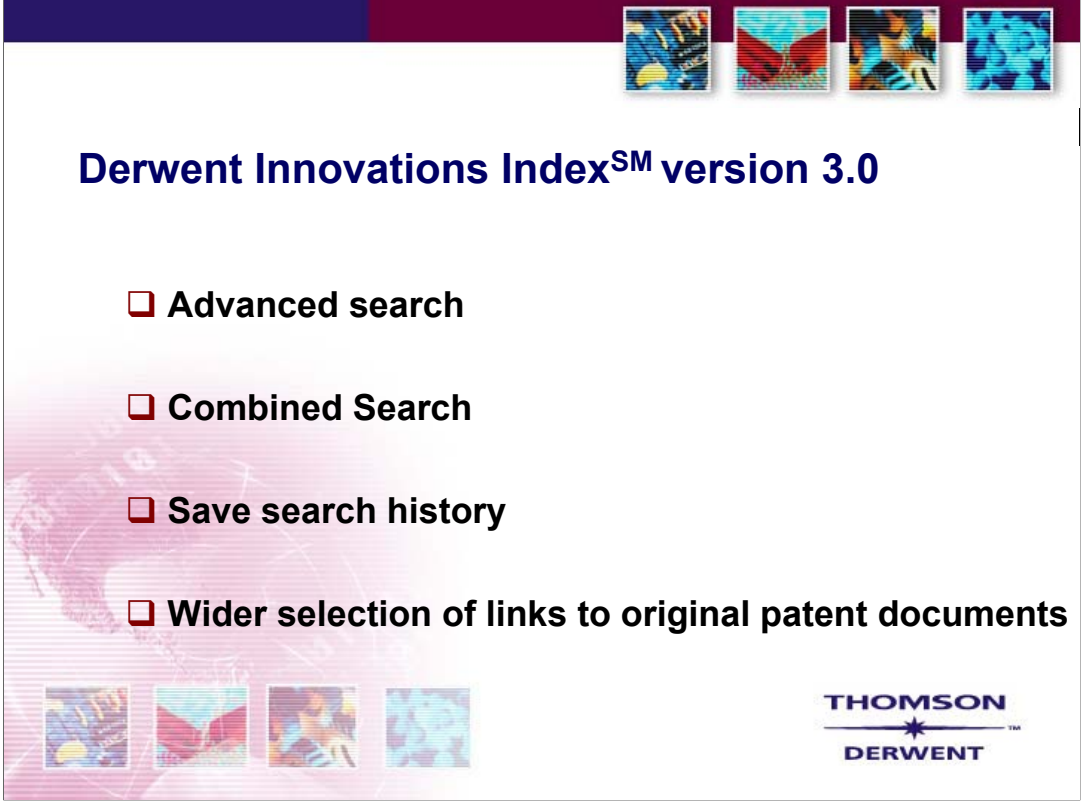
So far there is a difference in that Web services are Fixed-fee/unlimited use.

What's new on the web?

- Derwent Innovations Index℠ version 3.0
- Derwent Web of Software
- Derwent Biotechnology Resource℠
- GENESEQ.COM
- Derwent Discovery

THOMSON DERWENT

All of these are important, but I need to be selective



DII Combines Derwent World Patents Index and Patent Citation Index.

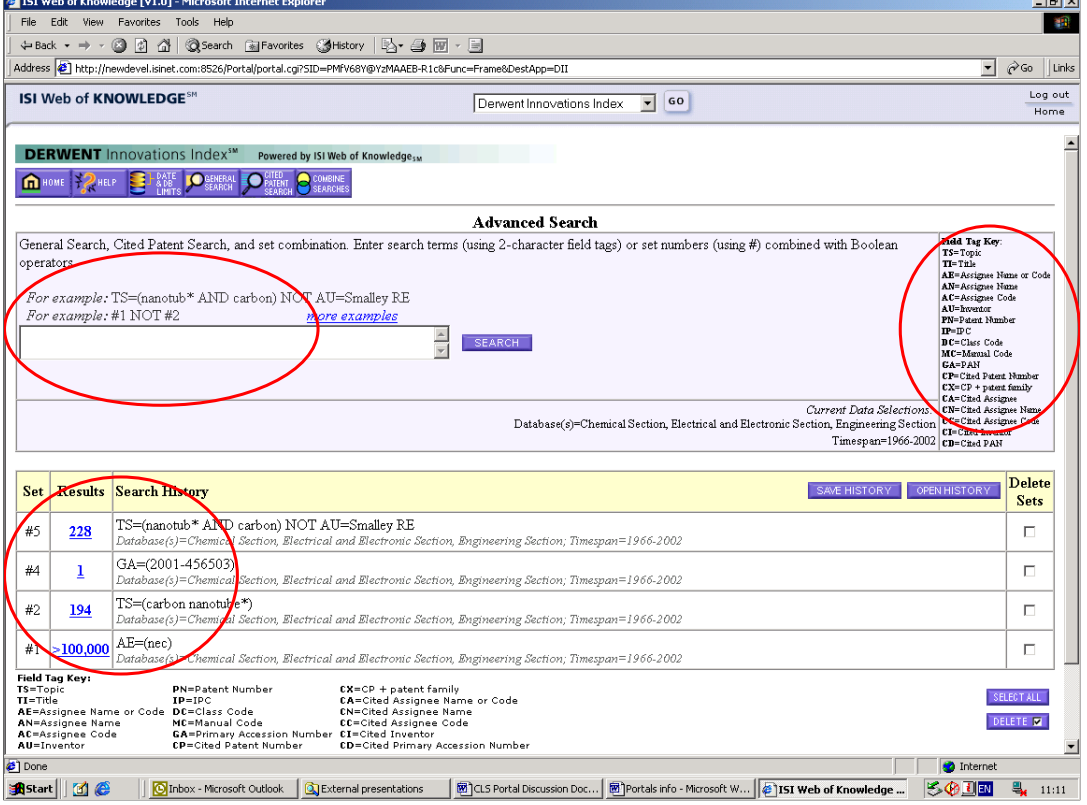
All the Derwent Records with the drawings are there

Patent Document Delivery in v3 has direct links to esp@cenet, MicroPatent and Delphion

It is fast and easy to use (very popular in the Britsh Library Patents Reading Room)

It links also to the non patent literature of ISI's Web of Knowledge.

Citation searching is an innovative added bonus.



This shows two of the key new features - the all important set-searching and advanced search option.

DII is probably the basis for many future developments. Not only can we expect the appearance to change, but a great deal of innovation is in hand for chemical searching.

A range of additional content is likely

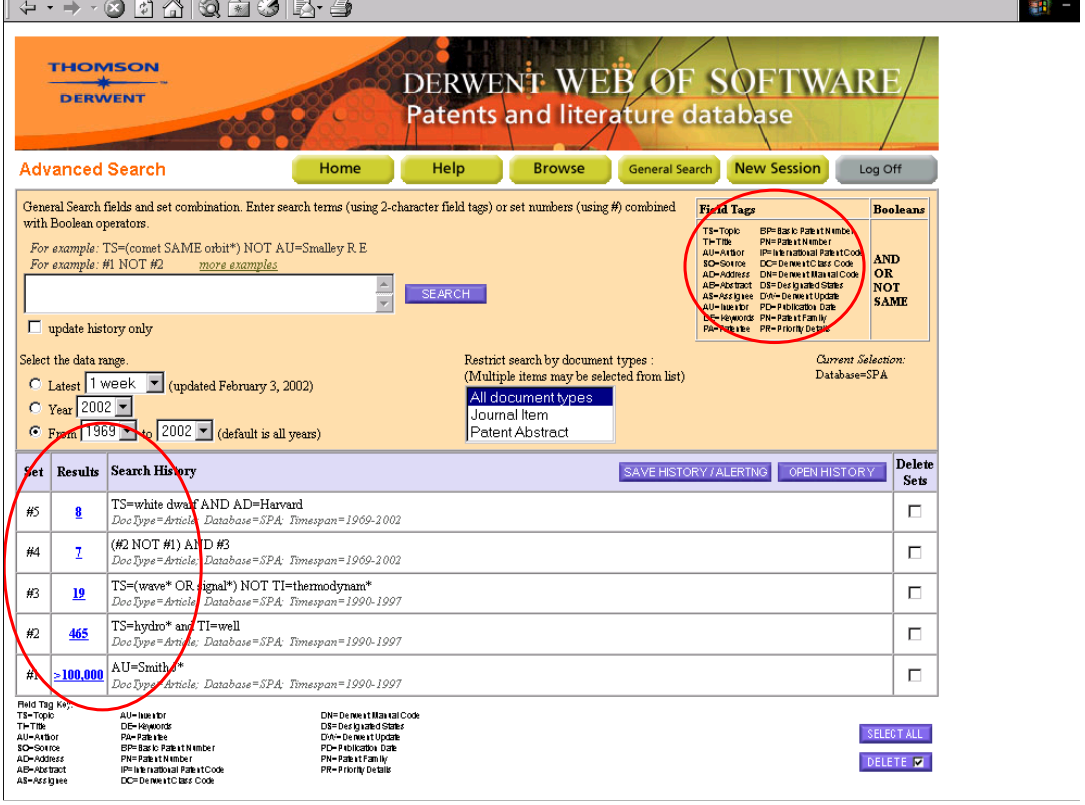
Derwent Web of Software

- Four software and e-commerce prior art sources available in a single cross searchable product
  - **Patents** – Derwent World Patents Index (T01 and T02)
  - **Literature and Conference Proceedings** – ISIs Compumath (800 titles)
  - **Gale group selected coverage** (1000 titles)
  - **Technical websites** (indexed by ISI)

THOMSON DERWENT

An area where prior art searching always demands looking way beyond possible previous patent documents is software.

We have put together a variety of sources for launch in September and will rapidly add more

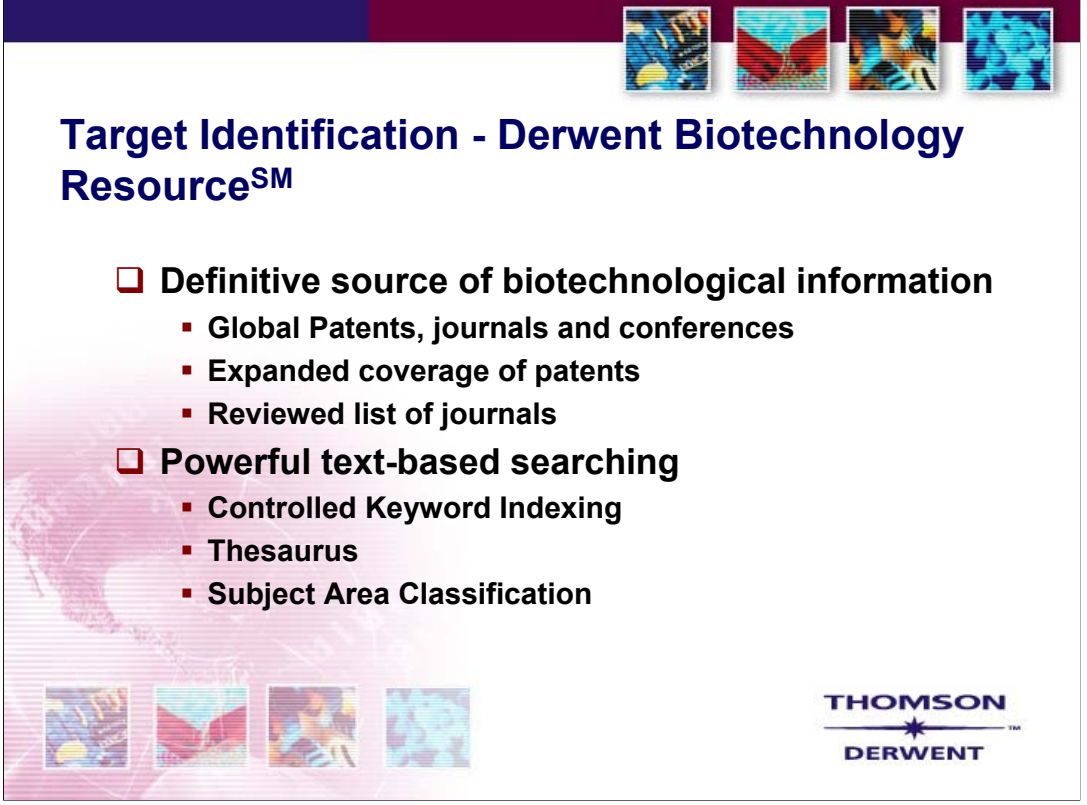


The advantage of an advanced search is the ability to save search histories and search between search data sets.

The field tags listing serves as an aid to boolean searching

Alerts can be set up weekly and run for six months, at which point they will need to be renewed. Alerts can be cancelled at any time.

The max. no of alerts that can be run at any one time set up at installation.



Derwent Biotechnology Resource is **The definitive source of biotechnological information.** It has its roots in our popular Derwent Biotechnology Abstracts, but is **much** expanded and designed for the Web

Derwent Biotechnology Resource contains a wealth of information for the pharmaceutical industry, including products,methodologies and new applications in the areas of:

•Bioinformatics (hardware, software and databases)

•Genomics and proteomics (including pharmacogenomics, toxicogenomics)

•Therapeutics (monoclonal atibodies, vaccines, gene therapy)

•Diagnostics (biochips, expression profiling and DNA probes/primers)

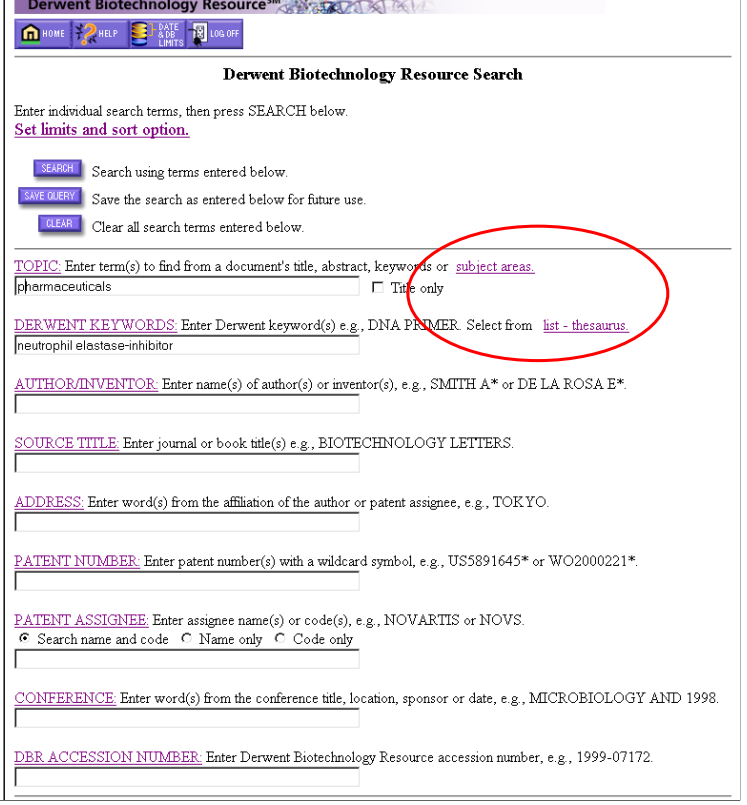
With global patent, journal and conference information in a single source, Derwent Biotechnology Resource offers powerful text-based searching which is vital for researching, in the case of company X, known drug targets which have been identified in GENESEQ.com.

Controlled keyword indexing terms are applied to every record in DBR by a subject expert. Also included are broader keyword terms which are also assigned. This is ideal where research is focussed on neutrophil elastase which is known to be a serine protein. You can search the narrow and/or broader terms using a controlled thesaurus to ensure highly accurate search results which are directly relevant to your research.

Target Identification - Derwent Biotechnology Resource

Subject area classification

Controlled keyword indexing



So company X chooses to search within DBR

Broad subject classifications are applied to each record. These can be selected from a controlled list. In this case, company X chooses the broad area of PHARMACEUTICALS which is entered into the TOPIC field.

Next, company X chooses to search for neutrophil elastase inhibitors using the thesaurus as a guide to ensure that the term is entered in the correct format.

The Web - in Brief

- An important focus for development
- Certainly attractive to end users
- Allows new approaches to handling indexing
- Lends itself to a portal approach
- XML standardisation of Derwent data adds even more possibilities

THOMSON DERWENT

What's Happening??

- **Derwent** part of Thomson since 1966
- **Derwent & ISI** have been moving closer for several years - logic in latest integration and customer benefits
- **DIALOG** was an important host for Thomson companies - needed to secure continuity for users
- **Wila** an important complement to Derwent & ***vice versa***.
  - Primary data
  - Wider markets
  - Synergy
- **Thomson** gives long term stability and growth

THOMSON
DERWENT

Derwent is not enormous - about 700 employees to cope with vast amounts of data and a global market, mainly of industrial customers. Many Patent Offices are larger - are very many of its customers

ISI, with a mainly academic market in literature is a little larger

Thomson comprises many specialist information companies in a variety of markets. It does not have any massive central Headquarters, but can provide investment and stability .

The sudden visibility stems from a move to a common corporate identity.


www.derwent.com