ELSEVIER

World Patent Information 28 (2006) 43–49

WORLD
PATENT
INFORMATION

www.elsevier.com/locate/worpatin

# Data visualisation tools—a perspective from the pharmaceutical industry

Jeanette Eldridge *

*AstraZeneca R&D Charnwood, Bakewell Road, Loughborough, Leicestershire LE11 5RH, UK*

This article has been developed from a presentation by the author at the International Patent Information Conference and Exhibition (IPI-ConfEx), in Benalmàdena, Spain, 6–10 March 2005.

**Abstract**

This article is based on some general overviews gained from my own and AstraZeneca colleagues' experiences of data visualisation tools and techniques, as applied to chemical and patent information, with some suggestions on how commercially available tools might be improved. Examples of the tools to be discussed include the visualisation components of programs used to process search results, particularly for chemical information and patents data, as well as specific packages targeted as mining and analysis tools, and also reminders of some of the simpler, but still effective, options available to assist in identifying patterns and relationships in data, even without integration with the more sophisticated programs now available. Comments are included on development suggestions and possible reasons for limited take up of these tools are offered.
© 2005 Elsevier Ltd. All rights reserved.

*Keywords:* Data visualisation tools; Chemical information; Patent information; Competitive intelligence; Intellectual property analysis

## 1. Introduction

The remit for the original presentation on which this article is based, was to provide a comparison of data visualisation tools, particularly for patent and chemical information. It is a recognised fact that representation of this type of information is challenging, particularly with the requirements to provide ever-larger answer sets for data and text mining, resulting in complex outputs and reports. Analysis generates a variety of visual report options, some of which may be easier for the ultimate recipient to interpret than others. Earlier publications [1,2] have provided a thorough overview of the tools available for patent information analysis, and newer developments have also been described in more detail than the current article has space for [3]. The aim of this paper is rather to bring together the assessments of a number of patent information specialists, who have experience of applying these tools to the day-to-day processing of patent and chemical information in AstraZeneca.

With this brief review of a range of packages, I do hope to highlight a number of new and established players in this arena, showing common themes and unique features, along with some identification of their main strengths and possible drawbacks.

The intention is not to provide an in-depth analysis of all the data visualisation tools available, as there is a wealth of software on the market, but hopefully the selection discussed here identifies some key players and types of visualisation used, and reflects that this is still a fast-moving area.

I also do not intend to discuss in any detail the concepts and tools for text and data mining per se—my focus is on some key issues in the consideration of data visualisation aspects, with a few highlights of particular tools.

Finally, I wish to discuss some ideas as to why some tools may not have been as successful as hoped in this area,

* Tel.: +44 (0) 1509 644402; fax: +44 (0) 1509 645569.
*E-mail address:* jeanette.eldridge@astrazeneca.com

0172-2190/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.wpi.2005.10.007

and especially to identify some opportunities for the future, either for specific tools or in a more generalised context.

The views expressed in this article are based on discussions with my information science and patent technologist colleagues within AstraZeneca, as well as on my own experience as an information professional using various tools to present chemical and patent information to colleagues and customers.

AstraZeneca is one of the world's leading global pharmaceutical companies, with over 64,000 employees worldwide, and eleven major research centres in seven countries. Our business is focused on turning ideas into innovative, effective medicines that make a real difference in important areas of healthcare.

Global IS&L are 130 dedicated staff located on eight R&D sites in North America, Sweden and the UK. We contribute to the achievement of AstraZeneca's business goals throughout the drug life cycle by applying our skills, expertise and scientific knowledge in the exploitation of published information.

## 2. Critical issues for evaluations

In reviewing visualisation tools, it is important to bear in mind the actual requirements that information professionals and our chemical researcher and patent specialist colleagues have—rather than to focus on how the tool is marketed. It may be unnecessary to spend a lot of time and effort processing and analyzing data, if the recipient would be happy to see the output of the raw data, or perhaps to simply review titles, keywords in context, or other straightforward sorting and grouping options already available quickly and at low cost.

One key element that has come up in my discussions several times, is the fact that visualisation should be about providing a view which is not easily available otherwise—this brings in the ideas around clustering and categorisation, something that is often a challenge to achieve "manually" on the volumes of information generated in these areas. If this new view on the information feeds into a better-informed decision-making process, this will be a key "selling point" for the product.

Some practical points will also influence the usability of the tools, such as any limitations to the range of data sets processed—there may be a relationship between a tool and only a restricted range of information types or resources—some tools are not designed or are hard to apply to additional data sets.

Of course, the aim is to enable the analytical tools and/or the analyzed information to be shared between information scientists and researchers, but there may be restrictions applied through the subscription and licence agreements which impact the desire to share these tools and their outputs, and prohibitive costs per user or for the initial download of datasets may influence their roll out within an organisation.

Another consideration is how to demonstrate cost effectiveness versus other methods already in use or other available tools, and how to show a return on investment against the time and cost inputs required to achieve an improvement in presentation of patent information, bearing in mind how challenging it is to quantify the true value of information.

Assessing the costs involved requires an evaluation of what investment might be needed in support, maintenance and roll-out of a new tool, particularly if it is something targeted at endusers rather than information professionals.

A further criterion to identify is how the package compares to other tools already established within the user population (we need to bear in mind the barriers which often exist to adoption of new tools and processes, where there are significant changes or retraining required), and how it might be integrated with other tools already in use, to improve our overall processes and decision-making—increasingly tools are required to fit with a series of packages exploited to maximise their potential, rather than to be a single standalone product.

## 3. Tools considered

As noted previously, the tools considered for this article are not intended as an exhaustive list, but are those with which there is some experience within AstraZeneca, and also some which appear to have some distinct differences or advantages.

A matrix of direct comparisons between these tools has not been at all easy to produce, so instead I have looked at each of the entries in turn for items to discuss, with some themes developed along the lines of their key features, similarities and unique offerings.

### *3.1. Online commands*

You may question why this entry has been included, but it is important to remember some of the simpler options can sometimes be the most effective—the human eye and brain are already pretty good at spotting patterns, so maybe a simply ranked, sorted or highlighted presentation of the information retrieved gives enough insight.

As an example, familiar online commands such as Analyze and Tabulate for STN allow further refinements in presentation of results into particular user-selected categories, for output to Excel, or perhaps further processed through the STN Express® Table and Report tools.

### *3.2. Excel*

It became apparent during the preparation of this article that Excel forms the basis of many of discussed packages' visualisation output options, and in its own right can be identified as a well-established visualisation tool.

Of course, an advantage of presenting information in this way is the relatively low cost, as well as existing user

familiarity with Excel charting and tabulation presentations, with no requirement for individual specialised visualisation packages.

There is also a number of plug-ins and add-ins for Excel which can be used to display chemical structure diagrams, and the program allows for further analysis and sorting of information to a variety of criteria, with the additional advantage of enabling data from a range of resources, both internal and external, to be merged—an example being patent data found from online searches combined with that from internally accessed patent gene sequence resources, used successfully within AstraZeneca to present collated data.

### 3.3. BizInt Smart Charts for Patents

The BizInt Smart Charts [4] tools have been long-standing favourites for me, and the Smart Charts for Patents package has become an established part of our repertoire at AstraZeneca for reporting patent and chemical information.

The tool has a lot of flexibility in application to a wide range of data resources, in terms of the number of host systems and data file formats that can be imported and manipulated.

Colleagues agree that the tool is particularly straightforward to use for presenting chemical patent information, and is good for producing simple, clear reports, allows comments to be added as an additional column, and key elements can be highlighted by, e.g., changes in ranking.

In Smart Charts, the retention of links to records and the preservation of sub-table layouts as retrieved, for example, from patent abstract databases on STN, is superior to Excel, and it is easier to display images, including structure diagrams, in Smart Charts and to update the charts on a regular basis with new data.

Data can be exported from a number of resources containing patent and chemical information, such as MicroPatent® and IDdb, as well as the online patent abstract databases, with easy combination of data elements into charts with nested sorting options. Simple sorting by common data elements, such as the basic patent number or priority date, allow easy grouping and duplicate removal for simple presentation of large amounts of patent information from these diverse resources.

### 3.4. VantagePoint

In 2004, BizInt Solutions introduced the option of integration with VantagePoint [5], which provides additional data visualisation possibilities. The integration with VantagePoint works both ways, allowing data to be exported from Smart Charts for analysis in VantagePoint, or data from VantagePoint can be exported to Smart Charts to create various reports.

VantagePoint itself provides a number of visualisation options, ranging from a matrix presentation of data, correlation maps with connections between data points, with drill down options to view the original records, and of course export to Excel for production of spreadsheets and charts. The package is applicable to patent outputs from a variety of hosts and databases, and the "import engine" is customisable to deal with data from other resources. Data clean-up tools are available, but as the volume of records grows, the subsequently generated maps can be very difficult to interpret, and users tend to resort to Excel outputs for simpler presentations.

Overall, VantagePoint's visualisation features are interesting, but this is one of a number of tools described in this article which have a relatively high cost per user, with an additional annual subscription which is not insignificant, especially as the intention is for the tool to sit with each user so the scope of the analyses can be explored—report outputs do not allow all of the analytical features to be shared.

### 3.5. Derwent Analytics$^{SM}$

When launched, Derwent Analytics$^{SM}$ [6] was exclusively based on Derwent World Patent Index data, but developments now allow data from Delphion to be processed.

The List clean-up option is an unique enhancement of the basic VantagePoint tool that also underlies Derwent Analytics$^{SM}$.

The various different maps generated give representations with increasing levels of complexity, so may need in-depth analysis to glean the key results.

I am told that the key to effective application of a tool of this type is to have a large number (3000–5000) of abstracts to analyze, so it is an advantage to have an "Open Access" agreement for effectively unlimited download of Derwent WPI data, but if this is not part of your company's information strategy, this can be a significant restriction, as the costs are likely to be prohibitive, because the inherent "fuzziness" of a large dataset means it is inevitable that a major proportion of the downloaded full records will ultimately be discarded.

### 3.6. MicroPatent®

MicroPatent® [7] worksheets are seen to have a useful range of output options, particularly for endusers, but it is felt that significant manual effort is often required to produce the desired MicroPatent® worksheets, so perhaps this feature is not so straightforward for the occasional user, inhibiting people from reaching the endpoint of visualisation of the data.

However, if the users persist, the export to Excel, with e.g. claims text included, is a popular option. The outputs can then be merged with data from other resources, such as records from in-house patent sequence searches, allowing enhanced reports with tables and charts to be produced.

As mentioned previously, the option to export to BizInt Smart Charts is also appreciated by information professionals, for the presentation of large amounts of data

in a simple format without the need for extensive preparation.

### 3.7. Aureka

Aureka [8], now an integral part of Thomson Scientific's MicroPatent® suite, has many in-depth analysis and visualisation features, and the product is strongly targeted towards IP portfolio management and analysis.

This is one of the products, which uses contour map displays for data visualisation, but in this sort of presentation, one must be clear about what conclusions are to be drawn from the position and proximity of data points, particularly as the peak labels may not be consistent or particularly helpful in deciding what the peaks actually represent.

IT problems made evaluation of this tool problematical in early 2004, so our licence was not pursued.

### 3.8. SciFinder®

SciFinder® [9] is a key tool, which has data visualisation capabilities, aimed at the end user, in contrast to some of the others described here, which are more likely to be used by information professionals.

Analyze and Refine commands enable data to be processed through to various categorisation options and displayed through the Panorama feature, which is similar to other cross-tabulation displays in other tools. Getting the best out of the Panorama tool is not always intuitive, but once grasped, can provide an excellent overview of the data presented. Especially attractive is the retention of links from the tabulated data to the original results and through to full-text.

Of course, full access to these valuable features comes at a not insignificant cost, if the product is provided to the desktop of all potential users under an annual subscription model.

### 3.9. STN Express®

STN Express® with *Discover!*™ Analysis Edition [10] has features for the information professional which resemble those available through SciFinder® for the enduser. For example, the Cross-Tab outputs are obvious parallel developments to the Panorama options seen in SciFinder®. So similarly to the SciFinder® Panorama output, we can display a cross-tab representation of the data, which has been previously grouped according to company name, and drill down from the totals cell to the records in question.

The Report and Table Tools are seen as particularly useful by patent information specialists, for reporting in a consistent format, which gives a common view to patent attorneys across the organisation, no matter at which site they are based.

The Data Grouping Tool is valuable to prepare the data set by automatically clustering similar terms, but this feature can also be customised to your own settings.

The STN Express® Analyze Plus Wizard allows the familiar ANALYZE and TABULATE commands to be used to analyze almost any field, and further refinements can be applied to the selection; it is particularly advantageous to be able to include multiple STN-hosted database outputs.

The Variable Group Analysis tool for R-group analysis is particularly attractive, and has obvious utility in summarising potentially extensive datasets in just the way a chemist requires. STN Express® appears to be one of the few tools, which provides this sort of analysis on chemical information from patents data from external databases. Of course, several structure–activity relationship analysis tools exist, and can be applied to chemical data from a variety of sources, but these do not appear to be routinely linked directly to external patent information sources.

Again, outputs in Excel chart format are relatively easily generated, which can then be captured for presentations or reports, useful for supplying to internal customers who require the summary rather than access to the detailed analysis process.

One downside is that, although we are able to use STN Express® as communication software to access other hosts like DataStar and Questel.Orbit, outputs from these hosts cannot be processed in quite the same way through the Tools, as the latter are optimised for STN datasets.

Once a data set is processed it is not easy to update, in contrast to BizInt Smart Charts, so to add data or to get a different view, we often have to start again from scratch.

A recent development is the integration of the latest version of STN Express® with a new visualisation package STN® AnaVist™, which is described in more detail in Section 3.15.

### 3.10. RefViz™

I wanted to mention RefViz™ [11] briefly, as its features do show another view on data which could be used on the textual and bibliographic elements of patent abstracts, although it is currently targeted at specific literature data sets; we have already found it to work with a wider range of data sets than those explicitly suggested by the suppliers.

With RefViz™, bibliographic records held within personal databases in Reference Manager, EndNote, or Procite can be analyzed and presented in two key views. The Galaxy View groups references conceptually, while the Matrix View grouped references according to terms discussed together.

The views are quite intuitive and give different insights on the data sets concerned to suit different user requirements.

### 3.11. OmniViz®

OmniViz® [12] forms the basis of the RefViz™ visualisation components. In its "native" form, the software can

integrate data from multiple sources, and has additional visualisation options and specific features directed to analysis. It was last evaluated some time ago within AstraZeneca, from the point of view of both information professionals and chemists, as it is able to analyze diverse data types and integrate scientific and patent data, but was not progressed due to the costs (per concurrent user and annually).

### 3.12. Vivísimo

Vivísimo [13] is a tool with potential to be applied to textual patent data, to perhaps provide simpler clustering and categorisation views on the data, rather than more advanced visualisation. Vivísimo technology also provides the advanced text clustering features in Aureka.

From their website, you can very quickly produce an hierarchical listing in ''real-time'', with named clusters and hit terms highlighted, linking through to original abstracts for easy review. It seems relatively simple for end-users to drill down within the descriptive folder names to the articles, and as this package can be applied to other datasets defined by the user, this could, in principle, be applied to patent abstracts or full-text documents.

### 3.13. anacubis™

When I put the original presentation together back in January 2005, I wanted to highlight anacubis™ [14], as I had seen a number of demonstrations which I found impressive, but I heard soon after that the parent company have withdrawn development of this tool within the patents arena [13].

A key application area was seen to be identifying the relationships between patents through citations and other connections, in order to explore licensing offerings—the intellectual property concerned is often several generations beyond the original patents, so it is important to know who really owns what.

So the sort of visualisation provided by anacubis™ is a little different from the others we have seen so far, and has potential in linking patent information with other business and company data—a shame if it disappears completely.

### 3.14. Other recent additions

Seen at IPI-ConfEx 2005, PatAnalyst [15] and PatentExaminer [16] are two products with similar profiles with regard to their interfaces, datasets and functionality, being based on the EPOQUE package used by EPO patent examiners, made available for commercial development as of last year.

Both have excellent visualisation and viewer functionality, with colour coding to reflect the frequency of term occurrences and their proximity to other terms—a really easy way to see highly relevant sections. These will need some usage to evaluate in comparison to each other and the existing range of tools available.

### 3.15. STN® AnaVist™

Most recently, STN® AnaVist™ [17], launched in July 2005, is a new visualisation component which works with STN Express®, to give information professionals a variety of ways to analyze and view information found in scientific literature and patents.

Capabilities available in STN® AnaVist™ include a workspace displaying a default of four data visualisations, dynamically integrated, including cluster and contour maps, histograms, and co-occurrence matrices. A key feature is the customisable data grouping feature to combine disparate entries for, e.g., company or inventor names. Our usage of this tool is limited so far but the product appears to have more flexible reporting options than some of the other tools previously reviewed.

## 4. Limitations to usage

I have already mentioned some aspects relating to cost, complexity, and data source restrictions, which may limit usage of one or more of these tools.

A further general point to make about these tools is that although they give a variety of excellent visualisation options, these may often be difficult to capture to reports or presentations, losing detail and interactive features like links back to full-text, redisplay according to different criteria, etc. So it may be that the tool itself needs to be shared, bringing with it issues of cost, maintenance, training and support.

Data visualisation tools are usually linked into data and text mining tools, so there is frequently a requirement for significant preparation of data sets and customisation of features before visualisation is successful; the knock-on effect of this is to require more information professional intervention.

It may be more from the information professional's point of view, but a better understanding of the underlying algorithms would be beneficial in appreciating the application of these tools—we or our IT, computational chemistry and informatics colleagues may identify in-house solutions, but these may have already been addressed by commercially available tools, which means the latter could be adapted to our needs rather than starting from scratch.

Despite the range of different tools available which provide a variety of ways to envisage patent information, it should be appreciated that they actually depend on a significantly smaller number of visualisation components, such as VantagePoint, OmniViz®, and particularly Excel. This narrower dependency could be seen as a limitation in itself, as there could be significant impact when developments start to diverge, different versions of the underlying package are required, and so on. One might ask what

contingencies are being put in place by the visualisation tool developers to minimise this risk.

## 5. Other issues impacting usage

Unanticipated file format alterations, such as changes to field names, copyright lines, etc., have an all too frequent impact on the functioning of programs such as the analysis and visualisation tools described in this article. What might seem to the database producer or host to be a trivial change in data elements or record formats, often has unanticipated effects on tools used to capture or process data. Examples are Thomson copyright line changes, affecting how records are captured in their own Derpict software, and record layout changes which affect imports to BizInt Smart Charts or reference management tools—so hosts and database producers need to work with these software tool developers to avoid this effect, which is irritating and time-consuming for us as customers to resolve each time!

I think perhaps we would be keen to see all of the more sophisticated visualisation tools described today to be applicable to more than the current limited number of data resources, particularly where a tool is very intimately reliant on the data resources from the same producer—and this will surely contribute to the longevity of the tool.

The philosophy behind information retrieval relating to visualisation is really undergoing a transformation at this time—we are told that information scientists really need to relearn *not* to narrow searches down, but to keep the data set selection as broad (and therefore large) as possible to allow the mining and visualisation algorithms to work to their best.

However, our experience is ''rubbish in''—''rubbish out''—so we want to combine mining and visualisation with refined answer sets, and *not* display the whole world of data—as I think was highlighted in detail by discussions at the IPIConfEx.

## 6. Development opportunities

So some ideas that have come out of my discussions with colleagues may already be represented in part or whole by tools already on the market, for instance backwards/forwards citation maps for all patents with citation information available, but in certain cases the data is not even available for all patents (e.g., for patents not from the larger patent authorities).

Things are of course improving, but some tools still focus on processing only subsets of data, e.g., US patents only, so CI analysis and insights are constrained and therefore the analysis is skewed.

Iterative or de novo processing of datasets should be easier to do, making different ''cuts'' of the data or refinements in the selections easy to display.

A challenge for data visualisation is to resolve the differences in the way patent family members are identified and captured across different resources, in order to produce as simple a picture as possible, particularly with reduction to patent families pre-clustering to remove duplication.

An ideal is to provide a way to identify relationships to build real competitive intelligence insights—not only between patents and drug development information, but also information such as conference presentations on NCEs, inventor/author affiliations, company alliances and so on.

Of course, ways to capture not only specific compound information for R-group analysis, but ways to capture biological data, and ways to compare and cluster similarities between generic representations would be a real boon.

## 7. Conclusions

So these few examples have hopefully helped to identify a number of useful tools and techniques for visualisation of patent and/or chemical information.

All of these tools enhance the analysis and interpretation of information by enabling visualisation through a variety of methods, and I have tried to identify some key strengths for each, as well as exploring the pros and cons as I see them.

Although I have also highlighted some possible limitations, these are not serious barriers in most cases, so I am sure these tools will continue to develop, and I hope some of the ideas suggested will be incorporated in those developments.

## Acknowledgements

I am very grateful for the help received from information scientist and patent technologist colleagues within AstraZeneca in the preparation of this article; their expertise and experience has been invaluable in identifying some key issues highlighted.

## References

[1] Trippe A. Patinformatics: tasks to tools. World Patent Inform 2003; 25(3):211–21.

[2] Rees P. Patent Analysis Software. Research Information; Autumn 2003. Available from: http://www.researchinformation.info/riaut03rees.html.

[3] Dürsteler JC. Patent Analysis. Inf@Vis 05/06/2005; Message Number 167. Available from: http://www.infovis.net/printMag.php?num=167&lang=2.

[4] http://www.bizcharts.com. Product information: info@bizcharts.com; technical support: support@bizcharts.com; contact address: BizInt Solutions, 650 North Costello Place, Orange, CA 92869, USA; telephone: +1 714 289 1000, +1 866 602 4184 (toll-free); fax: +1 714 744 1316.

[5] http://www.TheVantagePoint.com. Product information and technical support: VantagePoint@searchtech.com; contact address: Search Technology, Inc., 4960 Peachtree Industrial Blvd, Suite 230, Norcross, GA 30071-1580, USA; telephone: +1 770 441 1457; fax: +1 770 263 0802.

[6] http://thomsonderwent.com/products/dapt/derwentanalytics/. Product information and technical support: http://scientific.thomson.com; contact address: Thomson Scientific, Holbrook House, 14 Great

Queen Street, London WC2B 5DF, United Kingdom; telephone: +44 20 7344 2800; fax: +44 20 7344 2900.
[7] http://www.micropat.com. Product information: info@micropat.com; technical support: eu@micropat.com; contact address: Thomson Scientific, Holbrook House, 14 Great Queen Street, London WC2B 5DF, United Kingdom; telephone: +44 20 7344 2800; fax: +44 20 7344 2900. EU Customer Service: telephone: +49 89 18 93 36 77; fax: +49 89 18 93 36 99; eucustomerservice@micropat.com.
[8] http://www.aurigin.com/—see contact details listed for Micropatent.
[9] http://www.cas.org/SCIFINDER/scicover2.html. Product information and technical support: help@cas.org; contact address: Chemical Abstracts Service, 2540 Olentangy River Road, Columbus, OH 43202, USA; telephone: +1 800 753 4227 (North America); +1 614 447 3700 (worldwide); fax: +1 614 447 3751.
[10] http://www.stn-international.de/stninterfaces/stnexpress/stn_exp.html. Product information and technical support: helpdesk@fiz-karlsruhe.de; contact address: FIZ-Karlsruhe, STN International Europe, Help Desk, P.O. Box 2465, D-76012 Karlsruhe, Germany; telephone: +49 7247 808 555; fax: +49 7247 808 259.
[11] http://www.adeptscience.co.uk/products/refman/refviz/. Product information: http://www.refviz.com; technical support: ResearchSoft at +1 800 722 1227; agent contact address: Adept Scientific, Amor Way, Letchworth, Herts SG6 1ZA, United Kingdom; telephone: +44 1462 480055.
[12] http://www.omniviz.com. Product information: http://www.omniviz.com; technical support: support@omniviz.com; +1 866 792 3174 (toll-free North America); contact address: OmniViz, Inc., Two Clock Tower Place, Suite 600, Maynard, MA 01754, USA; telephone: +1 866 466 6484 (toll-free North America); +1 978 461 1250; fax: +1 978 461 1299.
[13] http://vivisimo.com. Product information: +1 866 294 8484; http://vivisimo.com/html/contact-sales; contact address: Vivísimo, Inc., 1710 Murray Avenue, Suite 300, Pittsburgh, PA 15217, USA; telephone: +1 412 422 2499; fax: +1 412 422 2495; Paris Office: telephone: +33 6 63 24 47 92; fax: +33 1 45 35 30 61. General information: questions@vivisimo.com.
[14] http://www.i2.co.uk/anacubis/. Product information and technical support: support@anacubis.com; contact address: info@i2.co.uk; The Visual Space, Capital Park, Fulbourn, Cambridge CB1 5XH, United Kingdom; telephone: +44 1223 728600; fax: +44 1223 728601.
[15] http://www.patanalyst.com. Product information: patanalyst@empolis.com; patanalyst@jouve.fr; contact addresses: http://www.empolis.com; empolis GmbH, An der Autobahn, 33311 Gütersloh, Germany; telephone: +49 5241 80 42906; fax: +49 5241 80 41825; http://www.jouve.fr; Jouve, 11, Boulevard Sébastopol, 75027 Paris Cedex 01, France; telephone: +33 1 4476 1492; fax: +33 1 4476 8637.
[16] http://www.questel.orbit.com/EN/Prodsandservices/patentexaminer.htm Product information and technical support: International Toll-Free HelpDesk: Europe: +8000 783 7835; USA: +1 800 456 7248; help@questel.orbit.com; contact address: QuestelOrbit, 4, rue des Colonnes, 75082 Paris Cedex 02, France; telephone: +33 1 55 04 51 00; fax: +33 1 55 04 52 01; clients@questel.fr; UK: uk@questel.orbit.com.
[17] http://www.cas.org/stnanavist/. Product information and service support for STN®, SciFinder®, etc. help@cas.org; contact address: CAS, 2540 Olentangy River Road, Columbus, OH 43202, USA; telephone: +1 800 753 4227 (North America); +1 614 447 3700 (worldwide); fax: +1 614 447 3751. For STN-related information in Europe: helpdesk@fiz-karlsruhe.de; contact address: FIZ-Karlsruhe, STN International Europe, Help Desk, P.O. Box 2465, D-76012 Karlsruhe, Germany; telephone: +49 7247 808 555; fax: +49 7247 808 259.

**Jeanette Eldridge** is currently a Group Leader and Consultant Information Specialist within Global Information Science and Libraries at AstraZeneca, with responsibilities towards the provision of an information service aligned to internal customer needs. Before joining AstraZeneca in 1998, experience was gained at Boots Pharmaceuticals, and later Knoll/BASF Pharma, in retrieving and analysing patent information, particularly in the context of chemical and pharmaceutical patent prior art. Previous positions have included abstracting and editorial roles with the Cambridge Crystallographic DataCentre, Elsevier Science Publishers and the Royal Society of Chemistry.