

EuroSpec: an international spectroscopic data base accessible for the scientific community

Nathalie Estephan, Douglas N. Rutledge

*Laboratoire de Chimie Analytique, Institut National Agronomique INA P-G, 16 rue Claude-Bernard, 75231 Paris Cedex 05, France
Tel: +33 1 44 08 16 60, fax: +33 1 44 08 16 53, email: estephan@inapg.inra.fr*

Introduction

EuroSpec is a EU funded project which started in January 2002. It develops an electronic database and archive for spectroscopic and other associated data from peer reviewed publications and other scientific literature.

The project will run initially for three years and develop electronic data submission protocols in close collaboration with the major scientific publishers and directly with authors.

This data base project addresses the massive deficit in the amount of quality reference spectroscopic data currently available to the analytical community when compared to the number of spectra acquired.

The largest electronic collections of analytical data represent 1% or less of the known chemical structures. The number of chemical substance registrations has now passed 35 million (table1). Nearly all of the spectra, which are recorded in industrial and academic laboratories each day, are discarded or are unavailable to the scientific community and even to those who acquired them.

Table 1. The number of known chemical substances Chemical Abstracts Registry
(<http://www.cas.org/cgi-bin/regreport.pl>)

Count	22,167,694 organic and inorganic substances
	34,262,812 sequences
Total	592465-25-3 chemical substance registrations

Partners

The EuroSpec Consortium comprises of :

- 1- ISAS, Institut für Spektrochemie und Angewandte Spektrochemie, Dortmund, Germany
- 2- Creon Lab Control AG, Frechen, Germany
- 3- INA PG, Institut National Agronomique Paris-Grignon, Paris, France
- 4- Universidad de Aveiro, Aveiro, Portugal
- 5- ICT, Institute of Chemical Technology, Prague, Czech Republic
- 6- LGC (North West) Ltd, Runcorn, United Kingdom
- 7- IM Publications, Chichester, United kingdom
- 8- SPECS and BioSPECS BV, Rijswijk, The Netherlands

EuroSpec project will establish the infrastructure necessary to produce high quality reference spectroscopic databases with links to available associated chemical knowledge.

The list of journals, which are currently being proposed by their publishers for inclusion in the start-up phase of the project, is long and includes important titles from publishers such as Elsevier, Springer, RSC, SAS etc.

At the end of the "EuroSpec" project, the rights to the data will be transferred to a charitable society which will use it for the good of the scientific community as a whole.

Data submission

There will be three main routes for spectral submission to the "EuroSpec" database :

- 1- From authors of peer-reviewed papers depositing data concurrent with submission of their paper to a journal,
- 2- From organisations with spectral collections who wish to donate these to the database,
- 3- From individuals who wish to submit spectra independent of any peer review process.

The scientific authors will submit their paper with the accompanying electronic spectroscopic data files to the database through a web interface. Figure 1 presents the workflow model for the submission of data.

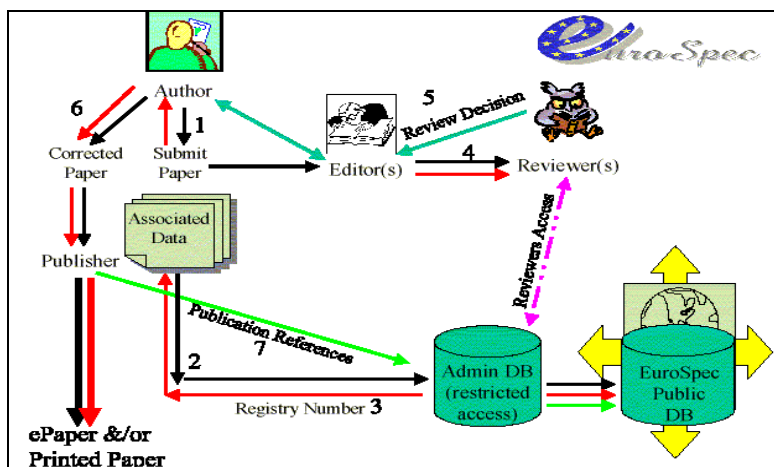


Figure 1. The workflow model for the submission of data

The data subscription can be done through this web site www.eurospec.org . You will be hosted by an interface introducing the Data Base and the European partners of the project to you (figure 2). The submission and the deposition of the data will be done through the IS-DB data entry system through a linkage which will be indicated to you: www.is-db.org

The data base will be powered by spectra and raw data which are associated to them. You will be guided throughout the deposition process.

EuroSpec accepts any kind of spectroscopic data (Infrared, NMR, Raman, etc.) and can support any file format.



Figure 2. The home page of the site www.eurospec.org

The International Spectroscopic Database (IS-DB)

The International Spectroscopic Database (IS-DB) went live on 19 May 2003. The system is capable of being used to deposit older data collections to keep them available for posterity. The files are always available in whatever is the current most widely used standard scientific data format.

Who to use the data entry system

The first step is to gain access to the Data Entry server by accessing the IS-DB homepage at www.is-db.org for registration.

Once logged in, a screen will appear welcoming registered users to the system and informing them of their registration number.

After registration, logged in users can then start depositing data in the archive, edit or modify information already on the system. If users wish to submit a data file, they may add it to a set of bibliographic information previously entered or they may start with new set of reference material.

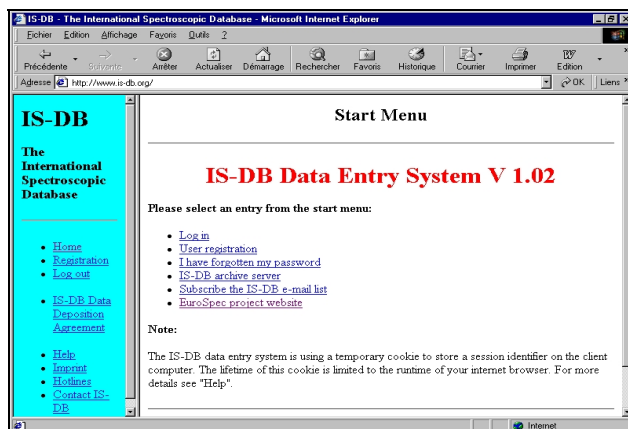


Figure 3. The IS-DB Data Entry system used to administer users.

With the data entered and the user logged out, the automation takes over. The relevant metadata entered are automatically transferred using XML data files from the Data Entry system to the Archive server. The Archive Server can be accessed either via www.is-db.org or through www.eurospec.org. The Archive Server is currently a customised version of the Q-DIS/PANDA electronic record management system.

The Archive can be searched by any of the keywords used during deposition, deposition date or a variety of other search fields.



Figure 4. The IS-DB Archive Server front page.

Conclusion

EuroSpec may be a new interface within the scientific community. It could provide a great exchange of knowledge and competence.

The stake of such a project exceeds considerations of a technical order. It is, in the long term, a question of changing mentalities within the world of research : to learn how to centralise dispersed knowledge with an aim of sharing them. Although financed by the European commission, this project interests the whole of the international scientific community, editors, authors or laboratories, which already understood the interest of collaborating.

This collecting of data:

- will support the exchanges of information and will authorise debates on an international level,
- will avoid double entries in terms of handling (and financial wasting which results from it),
- will reduce the delays taken because of a hard gateway to the existing data.

The existence of such collection of spectroscopic and substance data should stimulate significant progress in chemical research and have wide implications for human health, nutrition science, environmental protection and educational progress.

References

1. Supported by EU Grant No. G7RT – CT2001 – 05063 “EuroSpec – Access to Research Spectroscopic Data and Associated Chemical Knowledge”.
2. T. Davies, P. Lampen and P. Hughes, EuroSpec goes live at Achema, Spectroscopy Europe, Vol. 15 No. 3, p. 25-26, June/July 2003.