

Integrating INIS into a High Energy Physics Information Environment

Thoughts from CERN

J. Yeomans^a, T. Basaglia^a, R. Baudic^b, I. Picchioli^a

^a Scientific Information Group
ETT Department, CERN
Geneva, Switzerland

^b IUP DIST (Documentation et Information Scientifiques et Techniques)
Université Claude-Bernard
Lyon, France
E-mail of main author: Joanne.Yeomans@cern.ch

CERN, the European Organization for Nuclear Research, is the world's largest particle physics laboratory. Around 3000 people are employed by the institution and 6500 scientists regularly come from 500 universities around the world to use the facilities for their research.

Role of the CERN Library

The CERN Library must provide information to a huge number of users from a variety of subject backgrounds, countries, companies, and institutions, and not only whilst they are at CERN. These users might require material ranging across textbooks, research papers, standards or videos.

The Library aims to hold physical copies of all relevant monographs, conference proceedings and journals and also has a huge collection of grey literature, however, the main body of material is available online via the CERN Document Server (CDS). This database is not only the traditional library catalogue, but also the gateway to CERN's huge collection of online resources, integrating bibliographic records, online full-text and other online resources.

Increasingly, much of the material is received in electronic format and either linked or uploaded; printed material, especially grey literature, is also scanned to produce pdf or text files for uploading. Grey literature alone accounts for approximately 60,000 documents a year yet due to the automation of many processes, staff time is used more efficiently in development activities rather than procedural processes.

A semi-automated process is used for the uploading of bibliographic records from a variety of databases. In most cases this has been achieved by working with programmers to produce search and upload algorithms and has been implemented by special permission from the database owners. By this method we have been able to obtain basic records from some of the most useful high energy physics (HEP) sources: arXiv; Inspec; INIS; IN2P3; Mathdoc; and to use staff time more effectively to enhance these records, for example, with full-text links.

Typical CERN library users are very demanding and have similar needs and high expectations: often they are sophisticated searchers, they want to search quickly and can be impatient, they need to be able to locate specific documents when they need them, they want up-to-date information, and they expect to find most of the documents they need online.

The challenges for the Library, therefore, are to provide an integrated, yet easily navigable, information environment which is done primarily through the CDS software.

Identified INIS improvements

INIS has many good features but has some quite major drawbacks which deter our librarians from recommending it and HEP users from relying on it. It is worth noting that our environment prevents us from offering the database on CD so our users access it via the Web.

The main problem is the absence of links to full text: either to journal articles which are already available online, or to scanned NCL. This, combined with the lag between publication and input of records for important journals, degrades INIS's strength in grey literature. Our users are used to the one-stop-shop approach and expect to find grey literature alongside conventional literature, bibliographic information alongside full text.

Certain display features of the database cause frustrations for users: the default display order is opposite to their expectations, i.e. 'last in, first out', the display options are inflexible, the download functions limited. The direction of most major databases, especially for physics, is towards integration across sources and yet in order to find alliances, INIS will need to show features common to these databases. For nuclear physics we include those full-text databases commonly called preprint servers and it is to these that INIS must also look.

The database's Web interface can give unreliable search results and often times out altogether for complex search strings. These bugs give a lack of credibility to the data itself and must be rectified if users with a preference for Web are to be encouraged.

As the database grows, the accuracy of records needs to be addressed with an improvement of the use of the excellent thesaurus, and standardisation and clarification of resource-type definitions. Searching could be enhanced with the introduction of citation searching, and numeric data fields; these are the directions in which we see similar databases moving.

For the delivery of the INIS NCL to CERN users, Library staff are currently uploading full text from INIS on CD-ROM to the online CDS database. However, a way to convert pre-2002, multi-page TIFF files to PDF has had to be found as the TIFF format is virtually obsolete in our environment. For the kind of developments mentioned above it is problems like this that INIS will need to solve.

We believe partnerships with the following could improve the database's visibility and might help to drive the developments we hope to witness:

- (1) National Nuclear Data Center (<http://www.nndc.bnl.gov/index.jsp>).
- (2) Crossref – links to INIS alongside Inspec at online journal websites.

Conclusion

For CERN's users, we believe INIS should begin to focus on its Web product and the benefits and enhancements that Web technology can bring. Developments in line with other databases need to be considered in order to aim for integration of services. If this does not happen, INIS is at risk of becoming further sidelined by HEP users.

Automation of many processes can conserve staff time for future developments and CERN is keen to collaborate with INIS where this might be beneficial. We realize, however, that the user community for INIS is disparate and priorities for others might differ wildly from CERN's. A discussion of these varied needs is welcome in order to gain some consensus on the future of INIS.