

Report to the NRDC from the OECD Nuclear Energy Agency – Data Bank,**17 June 2003****INTRODUCTION**

The NEA Data Bank service statistics for 2002 continues to confirm the fact that these services are more and more solicited in Member countries. The computer program services, including the distribution of integral data sets, recorded close to 4700 requests, which is the highest figure ever. The increase in the request for computer programs was especially noticeable with a 35% increase from 2001.

The statistics from the nuclear data services also show an increased number of accesses in 2002, especially after the removal of the password restrictions in July 2002 for accessing the EXFOR and the Evaluated data. A very large increase in the volume of data retrieved in 2002 is the result of the release of JEFF-3.0 and of new versions of other evaluated data libraries.

The Management Board of the Thermo-chemical database (TDB) project decided in November 2002 to launch a new phase of the project. This new phase will be a 4-year project starting from January 2003 and will be devoted to the review of chemical thermodynamic data of Thorium (Th), Iron (Fe), Tin (Sn), and Molybdenum (Mo).

The Data Bank manpower situation was stable in 2002 and is expected to remain at the present level throughout 2003 and 2004. The A5 post as head of the Data Bank was announced in early autumn 2002, but the recruitment had to be delayed due to an imposed freeze of vacant posts, a result of the difficult budget discussions in OECD. The freeze was lifted in the beginning of April 2003. Two A2/3 posts, both allocated to the nuclear data services, are expected to become vacant, one in 2003 and one in 2004.

NUCLEAR DATA SERVICES**EXFOR/CINDA:**

The compilation of entries for the CINDA and EXFOR databases continues. For 2002 941 CINDA entries were distributed from area 2 (that of the Data Bank Member countries.) For this current year, 2003, many more entries are expected (~2000) as JAERI have kindly provided the CINDA entries relevant to numerous of the JENDL libraries (including JENDL-3.2, -3.3, -D/99) some of which were never compiled at the time. Further to this the JEFF-3.0 library, which was officially released in April 2002 will also be included in CINDA this year. Compilation of newly published work is also up-to-date for 2002.

The Data Bank has prepared the tables for the CINDA-2001 database entries and we are awaiting final decisions on the exact format for exchange before we finalise our loading/distribution programs etc.

Following discussions with the IAEA Nuclear Data Section, it has been agreed that the NEA Data Bank will assume the responsibility of producing and printing the CINDA book, including the database version on CD-ROM (CD-CINDA/JANIS). This agreement is due to the fact that most of the CINDA book customers are within the NEA Data Bank area and that it would be more economical for the NEA to produce the books in-house. However for this year's book the IAEA will provide the relevant Postscript/PDF files relevant for sending to the printers as the NEA have as yet not developed the relevant programs as we are waiting for the new format before starting this non-trivial task.

The Joint Evaluated Fission and Fusion File Project (JEFF):

Since the release of the JEFF-3.0 General Purpose Library in April 2002, the work has focused on the testing and the benchmarking of the file. An on-line form has been created on the Project's web site (<http://www.nea.fr/html/dbdata/jeff3feedback>) in order to facilitate the sharing of information among JEFF-3.0 users and developers. The form allows the users to describe their feedback (e.g. processing problems, data representation anomalies...). This information is then summarised and discussed at regular JEFF meetings. The agreed-upon corrections are made available for users who wish to use them before the next official release of the data library.

Contributions to the validation of the JEFF-3.0 library were presented by CEA/Cadarache and NRG/Petten at the JEFF meeting held on 28 April 2003. A series of fast reactor criticals (MASURCA, SNEAK) were calculated and the results show improvements with JEFF-3.0 compared to JEF-2.2. A series of ICSBEP benchmarks, including various combinations of fuel compositions and spectra, were analysed using various libraries, including JEFF-3.0, ENDF/B-VI.8 and JENDL-3.3. This study confirmed the results of a previous work presented at the US Cross Section Evaluation Working Group, namely that the most recent versions of the evaluated files lead to more consistent benchmark results. Validation efforts within the JEFF Project are aiming in particular at solving the problem of reactivity under-prediction for LWR lattices, an international study carried out in the framework of the NSC Working Party on International Nuclear Data Co-operation (WPEC).

An extension of the JEFF Project mandate for another 3 years was recently agreed. It was noted that:

- 1) The progress on the decay data and fission yields libraries is still slowed down by the lack of manpower.
- 2) There is an important decrease in the level of nuclear data activities in member countries. In particular, it is difficult for JEFF/SCG members to make commitments on a level of participation of their country in the JEFF Project for the coming years. Thus, the planning of activities and deliverables is uncertain.

The JANIS software:

The first official version of the nuclear data-plotting package JANIS was released in October 2001. Since then, about 700 copies of the CD-ROM have been distributed and important feedback has been accumulated, mainly originating from E. Dupont (CEA, Cadarache), A. Koning (NRG, Petten) and Christopher Dean (Serco Assurance, Winfrith).

A new version, which will address this feedback, is expected to be released in late summer 2003. It is planned that this new version would enable the user to access activation data described using files 8, 9 and 10 of the ENDF-6 format as well as NUBASE-formatted files. The "Computation" and "Weighting" features will be generalised to allow the user to combine several types of data (e.g. a combination of cross-sections and angular distributions for the production of differential cross-sections).

The structure of the data in JANIS-1.0 (serialised Java objects) is considered to be an important weakness since databases created with a given version of JANIS become obsolete when the data input routines are updated. To overcome this difficulty, a relational database technology will be used to store and retrieve the data. A Java-based database will be used in the next version to index ENDF-6 formatted files. The code will then directly access these files for the display and manipulation of the data.

It is also planned to incorporate the CINDA database into the next version of JANIS. A first version including CINDA is currently being tested. The longer term goal is to incorporate all the major databases containing nuclear data related information (CINDA, EXFOR and Evaluated data) in JANIS and to develop routines for updating these databases through on-line connection to the master databases stored at the NEA Data Bank.

Services to Nuclear Data Users:

The nuclear data services are to a very large extent provided through direct on-line access to the CINDA, EXFOR, EVA databases containing bibliographic, experimental and evaluated nuclear data respectively. Following the decision by the NSC Executive Group last year to eliminate the password protection for the database containing experimental and evaluated data, the Data Bank has noticed a significant increase in the number of real accesses. The term "real accesses" refers to the number of accesses after subtraction of Internet search engines, which constitute a substantial proportion of the total number of accesses. For more details about this and the number of accesses to the different databases, see the section below on "Computer Infrastructure and Development".

COMPUTING INFRASTRUCTURE AND DEVELOPMENT

Achievements in 2002

Hardware

Dual Internet Service Providers and Linkproof

One of our Internet Service Providers (ISP), KPNQWEST, went bankrupt in June 2002 and was replaced by the TELIA operator. By December 2002 the Internet links with Telia and Renater were both running at 2 Mbps. In the previous configuration, the traffic was being preferentially routed through only one ISP because of the artificial weightings being given in the Border Gateway Protocol (BGP). This problem no longer exists, as we have abandoned the BGP approach to one that relies on dynamically determining the optimum route for both incoming and outgoing traffic. The route balancing work is done by a device called "Linkproof".

Linux cluster architecture

A set of 4 Linux operating systems in rack mounted Dell equipment has replaced the Ex-Digital True-64 Intranet server. All common system files, Intranet files and user files are stored on a shared Network Attached Storage (NAS) device through a Network File System (NFS) protocol.

Cisco switches

The internal network equipment shared by the Data Bank and the Office Automation (OA) unit of the Agency was completely renewed. Fast switches have been installed in two different parts of the premises. The vital elements of the network have the ability to failover to their corresponding redundant partners.

Backup

An integrated file backup operation, servicing all of NEA, was implemented in spring 2002. All platforms (Unix, Linux, NAS and NT servers) have their file systems backed up to cassettes by the same unique tool. In addition the NAS uses daily and weekly snapshots to keep track of all file changes. This affords the capability to restore online any file lost within a month.

Online Services

Statistics of the usage of the Data Bank services rely on the logging of all accesses through the web site. It is necessary to eliminate from the logs various unrepresentative accesses such as 'Page not found' errors, images and icons that get downloaded with a page view and so on. It has become apparent in recent years that search engines account for a large fraction of accesses to the web site. For instance, the number of accesses to the Computer Program Abstracts was 690,000 in 2001 and 733,000 in 2002. The search engines alone accounted for 364,000 (53%) and 507,000 (69%) in 2001 and 2002 respectively. This year we have attempted to remove the search robots from the statistics. Although there still remain many cases where real end users have downloaded large parts of the site using freely available tools to accomplish this task easily.

When access to parts of the site is password protected, there is no interference from ‘commercial’ search engines. Such was the case in 2001 for the Computer Program requests and downloads and the EXFOR, Evaluated data files and JEFF documents. In July 2002, following a decision by the Executive Group, the password protection was removed from the nuclear data areas. As a result there was a rapid increase in the number of accesses to those pages.

Table 2. Number of accesses

| | 2002 | 2001 | 2001 |
|-----------------------|-----------------|---------------|-----------------|
| | Robots excluded | | robots included |
| Abstracts | 225299 | 322689 | 689953 |
| Program retrievals | 489* | 286* | 287* |
| Cinda | 2733 | 1916 | 2929 |
| EVA searches | 4400 | 2333^ | 2335^ |
| EVA downloads | 2915 | 1527^ | 1530^ |
| JEF documents | 12783* | 11104* | 20309* |
| EXFOR searches | 9063 | 6979^ | 7005^ |
| EXFOR downloads | 8872 | 7431^ | 7545^ |
| JANIS | 25068 | 10250 | 14397 |
| Web pages - CPS | 107491 | 102865 | 159881 |
| Web pages – DB | 27738 | 17007 | 22179 |
| Web pages - Data | 57138 | 54866 | 56839 |
| Web pages - TDB | 16997 | 16129 | 30268 |
| Total accesses | 487714 | 525722 | 971009 |

* password protected

^ Password protected until July 2002

Table 3. Number of Gigabytes

| | 2002 | 2001 | 2001 |
|----------------------|-----------------|-------------|-----------------|
| | Robots excluded | | robots included |
| Abstracts | 3.9 | 4.0 | 8.4 |
| Program retrievals | 9.1 | 3.9 | 3.9 |
| EVA downloads | 13.3 | 8.2 | 8.2 |
| EXFOR | 0.24 | 0.2 | 0.2 |
| JEF documents | 10.3 | 8.4 | 14.1 |
| Web pages - CPS | 5.7 | 2.2 | 4.1 |
| Web pages - Data | 2.3 | 2.3 | 3.3 |
| Total traffic | 44.8 | 29.2 | 42.2 |

The number of total ‘page accesses’ to the various services has not changed very much since last year and is at about half a million. On the other hand much larger files are

now being transferred through the Internet; in particular for the Computer program dispatches and the Evaluated files (as a result of the release of JEFF-3.0 and JENDL-3.3 in particular).

Current and recent work in 2003

Cinda and Janis

The Data Bank started to produce a CDROM version of CINDA in 2000. Since that time the CD-Rom database used was a version of Oracle which could only run on Microsoft Windows. We have now completed a version of CINDA based on the same Java database system used in Janis: Mackoi. At the same time, the opportunity was taken to integrate CINDA into the Janis environment. Janis now integrates all three nuclear data repositories managed by the Data Bank (Evaluations, EXFOR and CINDA) into a single multiplatform CD-based search and plot facility. The final Janis 2.0 release is expected in the summer.

E-mail spam and security

An enormous amount of spam E-mails arrive at the Data Bank and the increasing trend is alarming. It is also becoming more difficult to recognize which emails are spam, as, for instance, the emails could consist of just an image and a text analysis tool will be powerless. A highly recommended software system (called Mailsweeper) has been installed in February with the ability to intercept spam emails based on phrase lists which are updated automatically on a daily basis. The software also protects from viruses using a similar text analysis mechanism.

Plans for 2003-2004

Internet server

The current Internet server is a dual ex-Digital True-64 Unix system. It is expected to remain in service for the next two years. It is reliable and powerful enough to fulfil its main role as Web server. There are, however, signs that its configuration is becoming obsolete. A major operating system upgrade is overdue and there are already difficulties in finding or installing current versions of open source software (Apache-2, Java, Verity search, ...). It is therefore planned to install a standard Linux web server in such a way that Web accesses will be spread equally over the two servers (in a round robin set up). The Linux server will be able to fully replace the True-64 Unix server when taken offline for upgrades.

Parallel computing

It is planned to implement a code such as MCNP on the set of Linux machines currently used for the Intranet and other dedicated tasks. The 8 Linux computers at our disposal can be configured to run the parallel processing tasks during the night when normal activity is at a minimum. With this technique, it is expected that some benchmark calculations and computer program tests, running in tens of days, will be over in much less time.

ORACLE

A major upgrade of the internal ORACLE database management system software (version 8 to version 9i) is planned for this period. It will improve the operation of the

“hot standby” facility we implemented between the master ORACLE server and the backup server. Currently, the operation of failing over to the standby machine requires a good deal of manual intervention.

COMPUTER PROGRAM SERVICES

Acquisitions:

During 2002 in all 85 new or revised versions of computer codes were acquired and 103 were verified, tested and master-filed. Twenty-two new or revised compilations of integral experiments (SINBAD, IFPE and IRPhE) were acquired during 2002. The number of acquired and tested programs and integral sets of data exceeds the expected values. The increase of testing could be realised with the help of external consultants.

The trend in acquisition of packages from member countries and others participating in the computer program service is shown in the following figure:

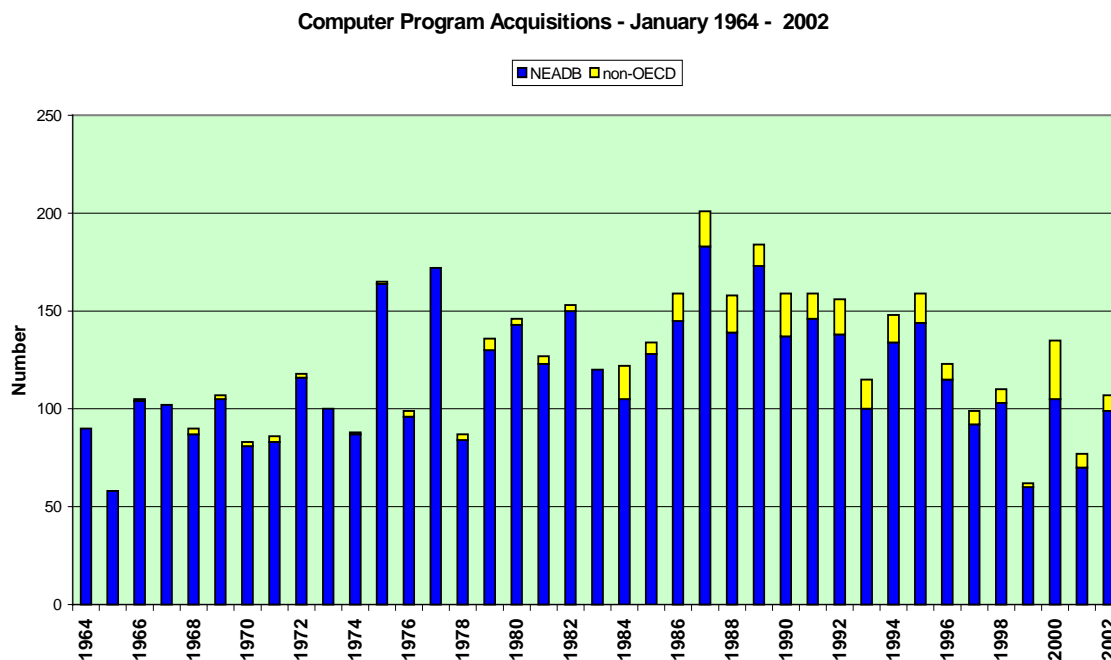


Figure 1. Total Annual Acquisitions of Codes and Integral Data

Dispatches:

During 2002, 2984 computer program packages were distributed, which is larger than expected. Regarding integral data experiments, 1680 sets were distributed during 2002, which is slightly lower than the figure of recent years. An additional 249 programs were distributed directly in Japan.

913 packages or 24% of the total distribution in 2002 were sent to non-OECD establishments or International Organisations.

A full historic for the past eight years is shown in Table 1 and Figure 2 shows the total dispatches (codes and integral data) since 1964 for OECD and non-OECD countries.

Table 1. Number of Computer Program Dispatches

| Year | Programs | Integral Experiments | Total | % Integral Experiments |
|----------------|-------------|----------------------|-------------|------------------------|
| 1995 | 1525 | 55 | 1580 | 3 |
| 1996 | 1641 | 208 | 1849 | 11 |
| 1997 | 1311 | 898 | 2409 | 37 |
| 1998 | 1787 | 1110 | 2897 | 38 |
| 1999 | 1898 | 2787 | 4685 | 59 |
| 2000 | 2249 | 1265 | 3514 | 36 |
| 2001 | 2205 | 1833 | 4038 | 45 |
| 2002 | 2984 | 1680 | 4664 | 36 |
| <i>Average</i> | <i>1950</i> | <i>1230</i> | <i>3205</i> | <i>38</i> |

NEA Data Bank Computer Program Services Annual Dispatches of Computer Codes and Integral Experiments 1964-2002

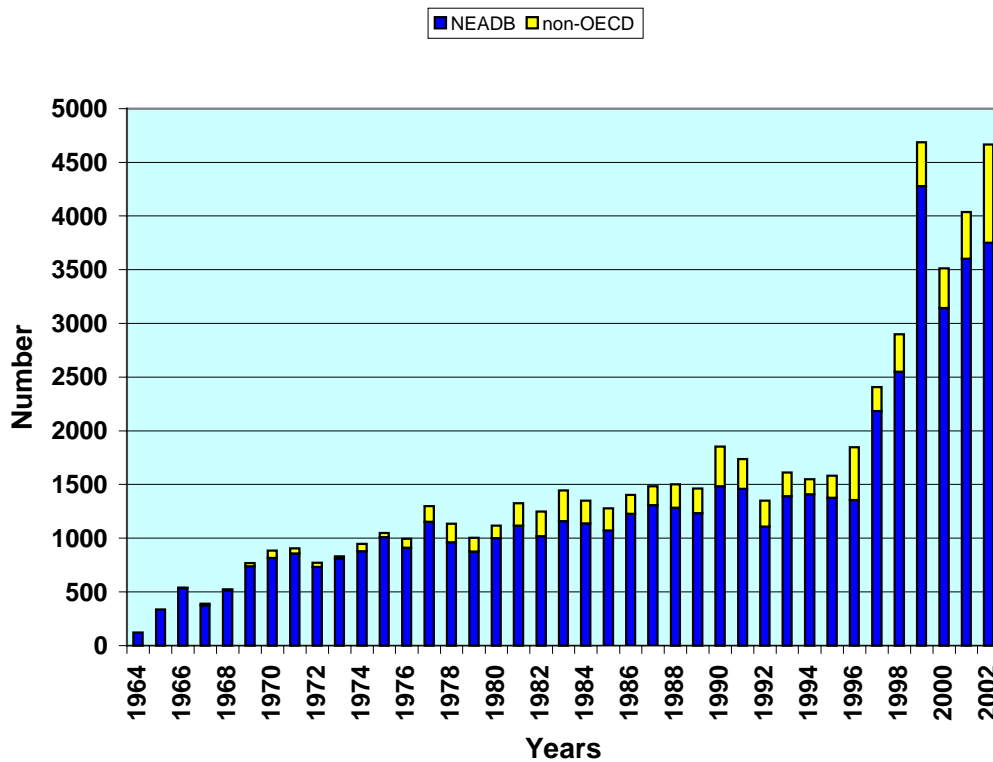


Figure 2. Total Annual Dispatches of Codes and Integral Data