INDC-185

PROGRAM AND PLANS

OF THE

NUCLEAR DATA UNIT

By the Staff

INDC-185

PROGRAM & PLANS OF THE NUCLEAR DATA UNIT

25 May 1967
Vienna

PROGRAM AND PLANS OF THE NUCLEAR DATA UNIT

I. Introduction and Present Status

At its meeting in Tokyo, September 1965, the INDSWG passed a resolution recommending that "the IAEA Nuclear Data Unit continue to explore on an experimental basis the possibility of acquiring, compiling and exchanging numerical data among interested IAEA member states and international organizations".

During the summer of 1966 the IAEA Nuclear Data Unit studied its position as one of the four world data centers, the others being Brookhaven, Saclay, and Obninsk. As the newest data center to be instituted, the IAEA Nuclear Data Unit was privileged to build on the experience of the other data centers, but was also obliged to seek a system of data storage and retrieval which would be compatible with those of the other three centers, as well as serve the compilation needs of a particular geographical area. A system was devised which was based upon those considerations. In designing this system, use was made of the experiences of the Sigma Center at Brookhaven, the Nuclear Data Compilation Centre at Saclay, and to a certain extent on the compilation experience of members of the Nuclear Data Unit. A scheme designated "Post Box", by which the proposed system could tentatively function to motivate international exchange on a systematic basis, was proposed in a meeting in which Dr. Zolotukhin represented the Obninsk data center and Dr. Colvin represented the Saclay centre. Regretably the American Center was not personally represented; however, since Saclay collaborates closely with Brookhaven, the American center was represented in a technical sense.

The IAEA neutron data compilation activity which is only part of the duties of the Nuclear Data Unit, was initiated at about the time of the Paris Conference on Nuclear Data, 17-21 October 1966. On this occasion the Nuclear Data Unit received a representative sample of current fission data from the major data producing countries of the world for purposes of comparing the various results. At the 25-29 October 1966 meeting of the newly named INDC (successor to INDSWG), the "Post Box" operation was tentatively approved, and most of the fission data that the IAEA had in temporary storage was released for purposes of international exchange. While the IAEA data storage and retrieval system serves a two-fold function, that is to say, as an interface for center-to-center exchange and as compiler of data for a designated geographical area, the Paris Conference principally served to give initial emphasis to data exchange. Since the fission data presented at Paris included some of the most recent fission results, it provided an incentive not only to the IAEA Nuclear Data Unit, but in a wider sense to the other centers (Saclay, Brookhaven, and Obninsk) to continue to have always available the very latest fission results. In short, sufficient fission data of recent origin is now in international circulation to warrant serious consideration of the promotion of this objective.

As INDSWG in the course of its few meetings became increasingly convinced that international data exchange was a goal that could be achieved, it also came to see the need for a more permanent committee to extend into the future the types of activity that it had performed on an ad hoc basis. The first step in this direction was taken in February 1966 with the formal directive by the Director General that the ad hoc INDSWG would be replaced by a more permanent INDC whose terms of reference would be worked out over a two-year period. These terms of reference have now been formulated with the aid of an expression of aims and procedures for the INDC, supplied by the committee itself, following a searching examination at its Vienna meeting, October 1966. A new stage has thus been reached in international cooperation in the exchange of basic neutron information.

This would be a logical time to examine the present achievements of the Nuclear Data Unit, and on the basis of its experience, to have a projection of its possible objectives and the resources that these would require. Perhaps it would also be useful to speculate about possible relationships between the new INDC, the IAEA in general and the Nuclear Data Unit in particular. This will be attempted, but it must be emphatically understood that these are complex questions which only future experience can answer; the only purpose here will be to invoke the present experience of the Nuclear Data Unit in order to invite further thinking on the subject.

In the discussions to follow, certain assumptions and definitions are tacitly taken for granted. Thus, it is assumed that neutron data compilation and exchange on a systematic world-wide scale is the commission of the four established centers already mentioned. Whether four are in reality needed is a matter that will be examined separately. A different but not completely independent question is that of "the criterion of success" which bears strongly upon objectives. Since compilation activity involves storage and retrieval of data, the implicit assumption is that data can be acquired, and if acquired, they can be stored and retrieved for the use of requestors. Thus, either one or both of two different criteria might be used as evidence of success, i.e., the number of data points stored, or the number of data points requested and retrieved. A much more severe criterion, and one which implies both criteria just mentioned, is first the extent to which a center or the system of centers can influence what is measured and how measurements are performed, and second, the extent to which a center or the system of centers can influence the numbers used for nuclear calculations and reactor design.

Since "evaluated data" is a term which frequently has a technical meaning, it needs to be defined as such and then used with care in order to avoid misunderstanding. When several measurements have been made on a given nuclide of the same quantity over the same energy region, it is frequently desirable to derive from these several measurements a "best set" of results. It may be that for a given nuclide several such energy regions exist. Especially in reactor applications it is not only desirable to obtain the "best sets" of results in the several energy regions over which measurements have been made, but it is also desirable to have the best theoretical interpolations between experimentally defined energy regions, as well as the most plausible extrapolation beyond. This is an undertaking for professional evaluators, and such data have come to be known as "evaluated data".

In the context of this writing and also in the program of the Nuclear Data Unit, "data" and "evaluated data" are carefully distinguished. It is recognized that the production of "evaluated data" depends first of all upon the efficient collection of all the most recent available data, and that to be effective, these evaluations must be performed by experienced staff. Hence, in general, neutron "data evaluation" can only be considered by the Nuclear Data Unit at some time in the future after data collection, storage and retrieval, as well as exchange, have been thoroughly established and are operating successfully. The first achievement of the Nuclear Data Unit, the evaluation of the experimental results pertaining to certain constants at 2200 m/sec, constitutes some exception to this general rule.

In the report which follows, emphasis is on questions of data compilation, storage and retrieval, and international exchange, and the closely related questions of long range prospects, i.e., the number of data centers, their relationships, the function of the IAEA Nuclear Data Unit, etc. Consideration is also given to other responsiblities of the Nuclear Data Unit, in order to provide a total picture of staffing requirements if these responsibilities continue to be carried in the future.

II. Nuclear Data Unit Activities

II.A. Data Compilation

Because of its general strategic situation, the IAEA was recommended as the site for a neutron data compilation center, charged specifically to:

- collect data from areas of the world uncommitted with respect to data collection
- stimulate international (i.e., center-to-center) data exchange.

After considerable study, the IAEA Nuclear Data Unit selected a data storage and retrieval system which seemed most suited for achieving this dual objective. This system had from its inception the following general characteristics:

- Data was to mean complete numerical results, which, by their character would be eligible for reference in CINDA;
- Data processing was to be devised so as to separate the operations pertinent to numerical data and bibliographical information;
- The index to the data was to be based on CINDA;
- Special consideration was to be given to the requirements implicit in the fact that the IAEA, for some time to come, would be functional as an interface between some of the centers;

- That the system be of benefit to the nuclear data producer, that is the experimenter, as well as the nuclear data user.

II.A.1. General description of the system

The data processing system, based upon the above listed characteristics, is described in some detail in Annex I. It can be said that, after the first six months of its operation, the system has proven to be more than adequate to fulfill the task at hand, and has received a positive reaction from our various correspondents from all over the world.

Data processed through the DASTAR-CINDU system undergo a number of several distinct stages; in order of their execution, they are:

- Requesting of data
- Acquisition of data
- Checking and correction of data, including proof-copy to the author (see Annex I)
- Dissemination of data

Each one of these stages has been organized into a logical sequence of steps, which take the data from requesting to dissemination with the most efficiency and reliability. Documentation is treated as part of the data to be stored and this part of the data entry must be thorough and complete as possible so as to include all information deemed essential to the user or desirable by the author.

Serious consideration has been given to deletion, correction and supersession of data. Data is withdrawn from storage and destroyed only in cases of some gross mistake in the data or in the case of transcription errors; a data entry which is revised on the basis of new information or re-evaluation, receives a label to indicate that the entry has been superseded, but it remains in storage. The arguments in favour of this approach are that: (a) it removes the possibility of accidental loss of data that should have been kept, (b) it reduces the amount of data handling, and (c) the evaluation of an experiment is preserved for the author or the evaluator. The probability that superseded data are used without knowledge is negligible because appropriate documentation is entered in both DASTAR and CINDU.

II.A.2. Rate of data processing and statistics

It is hard to estimate the real time involved in processing an average DASTAR data set from the step by step description of the data processing, as described in Annex I. Even if one assumed that a data center physicist devotes full-time to the task of data processing, it still does not take into account the delays inherent in a multi-personnel operation nor the dependence of the operation on the computer. However, data center physicists cannot devote all of their time to data processing: considerable system development is still taking place, data handling procedures are being evolved, and other duties and obligations take up considerable time. Important among the other duties just mentioned is the CINDA activity, which although coupled to some extent to the data processing, is nevertheless an independent activity requiring full-time attention.

One way to estimate the rate of data processing is to evaluate the progress made during a certain amount of time; thus, after approximately seven months of the IAEA Data Center operation, the following general stiatistics can be arrived at:

Data Processing

- Number of data lines initiated through the system: 114,000
- Of these, the number of data lines entered on the DASTAR master tape is 72,183. (The difference between these two figures reflects the time lag in the data acquisition process.)
- Total number of DASTAR data sets: 154
- Approximate number of data lines per DASTAR data set: 750
- Approximate processing rate 1 DASTAR/day per two physicists

The figure of 114,000 data lines was achieved by the BNL Sigma Center approximately two years ago.

Request Processing

- Total number of requests received and initiated: 190 (approximately 175 since the Paris Conference)
- Number of requests filled: 38 (20%)
- Number of requests pending (data not on hand): 152 (80%)
- Number of data lines sent out: 91,000

One figure of merit, worth considering at this point, is a comparison of the total number of data points in storage and the number of data points sent out on request. The ratio of the number of points requested to the number of data points in storage is 1.25 at the present time.

CINDA Processing

Contribution by the Nuclear Data Unit staff and readers abroad:

- From July 1966 to October 1966 22%
- From October 1966 to March 1967 18%

The decrease indicated for the second period corresponds to the increased data center activity and reflects the shortage of personnel.

II.B. Other Activities

Besides its function as a data center, which is number one in importance, the Nuclear Data Unit carries out other duties which will be listed as other compilation activities, meetings, standards, INDC secretariat, and Agency responsibilities. These will be briefly commented upon.

- Other compilation activities - the facilities list is a compilation which has been carried on jointly with ENEA. The possibility of maintaining a current (available but unpublished) compilation of this type should not be overlooked in view of its occasional usefulness and relatively small effort that is required to keep it up-to-date.
- Meetings - experience shows that two technical meetings per year will absorb the entire time of one physicist.
- Standards - this activity has essentially occupied time only under the category of meetings and hence has consumed about ½ man. Under this heading it is reasonable to include monitoring of the 2200 m/sec constants.
- INDC secretariat - a fraction of a man/year.
- Agency responsibilities - about a half man/year.

II.C. Staff Situation: present status

The following minimum staff requirements are deemed essential for the Agency's activity in the field of neutron data compilation and international exchange:

- 3 full-time P-3 physicists
- 1 full-time programmer
- 1 half-time secretary for book-keeping, correspondence, etc.
- 2 half-time key-punch operators

The present CINDA operation at the Agency requires:

- 1 full-time P-2 physicist
- 1 part-time programmer
- 1 part-time secretary for book-keeping, correspondence, etc.
- 2 part-time key-punch operators.

Under the present commitment of the Agency to participate in the world-wide CINDA operation, it is anticipated that the above staff requirements will remain constant. Thus the total present minimum need for data compilation (including CINDA) is:

- 3 P-3 physicists
- 1 P-2 physicist
- 1½ programmers
- 1 secretary
- 2 key-punch operators

It must be pointed out that the Nuclear Data Unit leader and his secretary have not been included since they are not operational in the CINDA and Neutron Data Compilation activities, nor the necessary professional staff for meetings, standards and INDC business (see II.B.).

By comparison, the Neutron Data Compilation Centre of the ENEA has a staff of 18 people, 6 of which are physicists; in addition, they have three permanent long-term consultant positions, which are filled by physicists and computation experts.

In contrast to the level of activity needs, mentioned above, there are available now for compilation and CINDA:

- 2 full-time physicists (H.D. Lemmel and A. Lorenz only, since K. Ekberg's and P. Otstavnov's time is taken up by conference and INDC secretarial work.) The present minimum requirement is four full-time physicists.

- 1 part-time senior programmer (P.M. Attree), and

- 2 part-time junior programmers (J. Chelmick and F. Hirschbichler). All 3 programmers spend a great deal of their time on work for other sections in the department; thus the minimum requirement of 1½ full-time programmers is at the present time not fulfilled.
- barely sufficient secretarial staff and key-punch operators.

Although great strides in neutron data compilation have been made during the initial stage of the Nuclear Data Unit center operation, it was accomplished with a minimum of staff. It should be pointed out that in order to have a fully efficient and complete operation, a larger staff, devoted specifically to data compilation, is needed. Unfortunately, the present financial situation at the IAEA is such that additional staff cannot be hired in the near future. In light of this limitation, careful consideration must be given to the commitments of the Nuclear Data Unit and to the judicious use of its present staff.

III. The Future Program of the Nuclear Data Unit

The long range future of the Nuclear Data Unit will certainly depend upon immediate successes in its present activities. According to views expressed by the S.A.C.*, the most important activity of the Nuclear Data Unit is data compilation and exchange; the success of this activity is to be judged not merely by the statistics of data compilation and exchange but by manifestations of these activities as well. Thus while it may be possible to consider long range plans of the Nuclear Data Unit, especially in the framework of the four data centers (Sec.IV) it is essential first of all to consider the plans by which immediate success is to be achieved.

III.A. Data Storage and Retrieval

The most important manifestations of data centers activities are likely to be rather far in the future. It is necessary therefore to establish some short range objectives by which progress can be judged. Compilation centers can be justified in the view of many, only if they can perform some useful function which is essentially unattainable by the time-honored publication methods. Modern possibilities for

* Scientific Advisory Committee

computer storage and retrieval together with efficient use of available communication techniques can in principle provide an access to all current results, which are unattainable by traditional publication means. Ready access to all current results is of course the purpose of the data centers, but this purpose is sometimes obscured by restrictive procedures. Success requires at least two steps: (a) proving that a suitable data handling system, or systems, can be devised, and (b) establishing user acceptance of such a system or systems.

We have already stated that the data center at IAEA has a dual function to compile data and to stimulate international data exchange; we also asserted that the principal immediate user is the reactor physicist or evaluator although the principal user in the long run might be the experimenter himself. As a result of the Paris Conference, an emphasis was placed on Cross sections of Fissile Nuclides by which a most recent set of such results were placed in circulation; in this achievement the possibility of effective center-to-center collaboration was also demonstrated. In view of the foregoing remarks, the Nuclear Data Unit intends in the immediate future to continue to give first priority to data on fissile nuclides. The immediate objective would obviously be to make internationally available data from experiments of distinct value to reactor physicists and designers.

Because the volume of fission data that has been processed up until now has used the present Nuclear Data Unit staff to its capacity, not much effort has yet been placed on the cross sections of the non-fissile nuclides. Nevertheless, plans for the immediate future include an increasing emphasis upon non-fissile nuclides and an appeal for increased experimenter participation in the geographical areas generally associated with the IAEA in the area of "data" compilation and exchange.

In zealously attempting to divorce data evaluation from data compilation, there is danger in overlooking the fact that evaluated data are like experimental results, viz., there may be several different "results" and these ("evaluations") like experimental results can have immense value, provided those which are current can be distinguished. The Nuclear Data Unit is of the opinion that "evaluated data" like "experimental data" becomes the object of international exchange. In this connection, we would like to acknowledge the receipt of a sample of evaluated data from Dr. Colvin of the NDCC at Saclay, which could serve as an initiation of evaluated data exchange through Operation Post Box, in a way similar to that in which the basic neutron data is handled. It should be added that this can be done without confusing the two types of data.

We conclude this discussion with some remarks pertaining to a slightly more distant future. At the present time, the neutron data activity of the Nuclear Data Unit is strictly limited to compilation. Compiled data, as treated at the Nuclear Data Unit, is in a very real sense the exact opposite of "evaluated" data; in the former, data never lose their original identity whereas in the latter, they lose all original identity. Between these two extremes lies an area of data which will be of greater value to most users than the compiled data alone. To avoid confusion with "evaluated" data we will term such intermediate data as "derived" data. "Derived data" is a term which means that the original data has in some way been transformed, e.g., it may be re-normalized, it may be merged, the variables may have been changed, etc. (never, obviously, is the compiled original data removed from storage). This type of supplementary activity (which is not to be confused with producing "evaluated" data) requires the judgement of experienced physicists. As soon as the position of the Nuclear Data Unit as a principal data center becomes consolidated, the present activity should be expanded to include producing (and compiling) this "derived" data.

In the view of the present Nuclear Data Unit staff, the anticipation of such an expansion means looking only a moderate distance into the future; it would call for perhaps one more physicist, a programmer and some additional auxiliary staff.

III.B. Other Activities

If the Nuclear Data Unit is to continue, without major increases in personnel, to promote meetings other than those which are consultative in character (i.e., if it is to conduct panels, seminars, conferences, etc.), then it is essential that a minimum of effort be consumed by "in-house" procedures and a maximum amount of effort be devoted to "meetings". According to present procedures for establishing meetings, approximately two years are required to guarantee a successful culmination. Thus a topic which matures quickly, or has been generally overlooked, has little prospect of being included for a meeting, excepting at a considerable cost in time. Considering the two-year "turn-over" time of key persons, the negative effect of inadequate long range planning is evident. A possible solution to this situation, so far as nuclear data is concerned, is to request certain "blocks-of-time" for certain classes of meetings. Thus, for example, a meeting of the "Paris Conference" type might legitimately be anticipated roughly every three years; a panel on standards might be anticipated every four years, etc. This procedure is to be recommended on the grounds that, at least at present, a meeting can easily be dropped or substituted, it cannot easily be added. A case in point is that of a conference or seminar on Fast Radiative Capture which Sweden would have co-sponsored if the meeting could have taken place in 1968. However, already by December 1966 such a meeting could already be ruled out. According to the S.A.C., meetings of major proportions, i.e., symposia, conferences, and possibly panels, should not be the concern of the Nuclear Data Unit. Thus, in order to promote international nuclear data meetings, it may be necessary for the INDC to cultivate direct contact with the Physics Group of the IAEA, which could be instrumental in the promotion of such meetings.

Because the IAEA has suitable facilities for conducting compilation activities, there are two activities in addition to neutron data compilation which can be contemplated with modest personnel support. One is continuation of the facilities list on a basis like the data compilation. The facilities compilation, as visualized, would be one in which the facilities information already on hand would be subjected to annual or bi-annual up-date, without publication; the results would be available only on request. Another compilation has been suggested by the recent Panel on Nuclear Data Standards which recognized the stimulation to standards activities that would result from a compilation of neutron standards activities as well as available services and materials.

Other activities have been carried on by the Nuclear Data Unit in the past. These include an evaluation of the 2200 m/sec fission constants and secretariat services for the INDC. At its recent Panel on Standards for Nuclear Measurements, the panel members looked on the

2200 m/sec evaluation as a standards activity and advocated extending or broadening this type of activity. Such an evaluation activity would certainly absorb the time and effort of one physicist. Looked on as a standard activity, a program of evaluation of reactions of special value to neutron cross section measurements is clearly distinguishable from "data evaluation" in the usual sense. Of course, an alternative is to regard the whole subject in the "evaluation sense" in which case the same objectives could be attained although it would clearly be desirable to treat the "standards reactions" as a unit in this case, and promote the work being done by evaluation groups of individuals.

IV. The International Situation

This report has concerned itself so far with factual statements about the activities of the Nuclear Data Unit, as they presently exist or can be extrapolated to the near future. If costly false moves are to be avoided in the future, it is essential to attempt to get some ideas of what lies ahead. Among the initial problems of establishing a neutron data center at IAEA were those of finding what kind of activity would best benefit the system of data centers and ultimately data users throughout the world.

It is not possible to consider the long-term objectives of the IAEA data center without making certain assumptions which have direct interaction with the activities of the other data centers.

The present situation is that a fruitful geographical distribution of labour has been established such that the BNL Sigma Center is responsible for the USA and Canada, the ENEA Neutron Data Compilation Centre, Saclay, for the OECD countries of West Europe and Japan, the Informazionnyj Zentr po Jadernym Dannym (Nuclear Data Information Center), Obninsk, for the USSR, and the IAEA Nuclear Data Unit for all other countries in East Europe, Asia, Australia, Africa and Latin America. This means that each of the four centers is to collect the neutron data from its geographical area, offers its services and answers requests to individuals in its area, always in close contact with and exchanging data with the other centers.

The first most obvious question concerns the required number of data centers. Our assumption is that four basic data centers is about optimum, for the following reasons:

- 4 centers provide adequate coverage of world's produced data, and each center is small enough as to offer individual services.
- Each part of the world covered by the centers has its own characteristics which require different procedures and methods.
- We have experienced that the existence of 4 centers provides a fruitful competition of ideas with mutual benefits.

Compilation can only be useful if it is complete; that means that each center has to fulfill its obligation within its area defined by the geographical distribution of labour. In this respect, we realize that the holdings of BNL and Saclay are by far more up-to-date and complete than those of the Obninsk and IAEA centers. It is therefore our present obligation to collect the data from our geographical area, the old as well as the new data.

The principal argument for consolidating the number of centers at the present time would be to achieve greater economy. Against this is the argument of loss of flexibility. In particular, each data center has developed certain techniques for handling data collection in its particular geographical area of operation, and these would be lost if consolidation were complete, with probable loss in effectiveness in data collection.

Assuming four data centers, what are the alternative relationships for most effective international data exchange ? The present situation as defined by the above mentioned distribution of labour, should be summarized first, and for this and the following discussions it is convenient to employ a graphical representation:

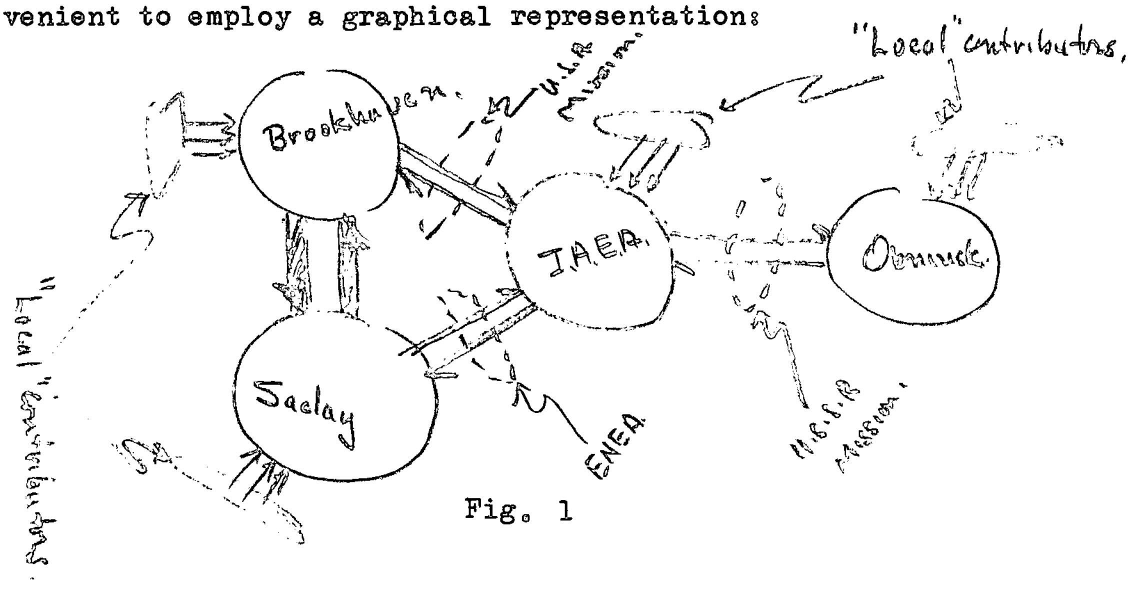

Fig. 1

Figure 1 is intended to portray the manner in which the IAEA data center is at present serving as an interface between the center at Obninsk and the centers at Brookhaven and Saclay which are relatively highly coordinated to each other. The figure is intended to also portray (a) the existence of a relatively large traffic of data between Saclay and Brookhaven (b) that local contributors have access to data-at-large through the data center with which each is affiliated. Rough agreements were accepted at the October 1966 meeting of the INDC. These were: (a) operation "post-box" which was the system wherey data for international exchange was voluntarily (assisted in some cases by requests from the Nuclear Data Unit) sent to the Nuclear Data Unit which then acted as a dispatcher to perform any necessary interface conversions and transmit to the other centers; a part of operation "post-box" is the periodical issue of a catalogue of data holdings; (b) channels for requesting data were to be as shown in Fig. 1. Fig. 1, which assumes four equivalent data centers, continuing to operate subject to the present interface incompatibilities, with the IAEA Unit continuing to function as an interface while also acting as a center for data collection from designated areas, is one possibility for future functioning of the data centers.

A completely symmetrical situation is indicated in Fig. 2:

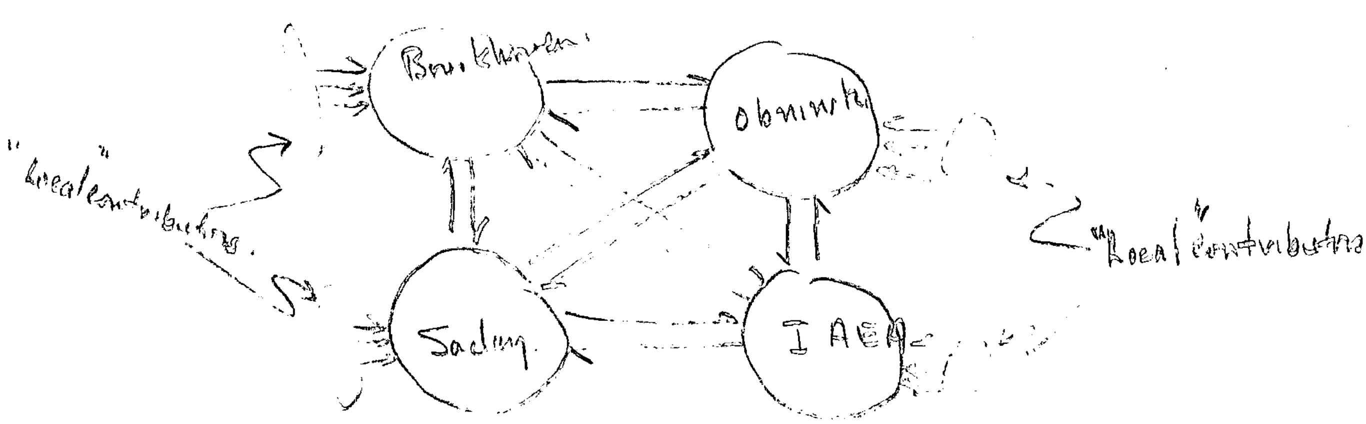

Fig. 2

Such a system is obviously remote at present. However, the situation in Fig. 1 could rather easily develop gradually into that shown in Fig. 2. The main reason for a cooperation according to Fig. 1, are the presently existing technical problems of format conversion which are solved by the IAEA center by converting the Obninsk type punched cards. In future, these technical problems might become less and less important, and the cooperation might change into the symmetrical system of Fig. 2. In this case, the IAEA Nuclear Data Unit would cease to act as an interface and would restrict itself to serve its own geographical area, thus saving some manpower for reviewing work which becomes more and more important as files grow.

The purpose of the preceding schematic discussion has been to vindicate the statement that in the long term, the data handling techniques and especially the staff requirements of the IAEA data center will depend upon the future complexions of the other three centers; the converse, of course, could also be argued, viz.: the complexion of the IAEA center will in the long run influence future complexions of the other centers. The problems of center-to-center exchange are probably not yet well enough defined to attempt any analytical decision on the long term relationships between centers, although such a decision will ultimately be needed. In the meantime, the Nuclear Data Unit proposes to continue to function as it has been doing from the relative position shown in Fig. 1.

Regardless of how the linkage between the centers works, there is the completely different question of which data should be held by which center. There are two extreme possibilities:

- All 4 centers hold identical master tapes of the world's data
- Each center holds generally only the data of its geographical area and asks for the data of the other centers only if requested.

The answer to this question depends completely on the frequency of data exchange and data requests, roughly on the ratio of the number of requested data to the number of stored data. If there are very many data requests from all over the world, it is of course advisable that complete master tapes are held by all centers. However, the difficulties involved with updating four identical master tapes should not be underestimated; identical master tapes is an "ultimate" which cannot be achieved on a four-center basis until many other difficulties are overcome. Thus we have to envisage the four centers as constituting a system in which each center holds the data of its geographical area which it makes available to the other centers on request. If all centers do not have all data then, it is a condition sine qua non, that each center distributes periodically a catalogue of its holdings. From the beginning of the development of our DASTAR system, we saw the need of having a catalogue of our data holdings. The result was CINDU, which was developed from CINDA, and which has been generally well received. Saclay recently issued a catalogue of its data files in Newsletter No. 4; since the amount of this data is steadily increasing it can be inferred that future Newsletters will contain catalogue up-dates. Similar catalogues from the Sigma Center and the Obninsk Center would be most useful. It should be added in conclusion that CINDA now contains an entry (Column 80) to verify if, and to state where, the data that is referenced can be obtained. This is a first step to provide CINDA with some additional features, just as CINDU has done on a more extensive scale. Some generally accepted modification of CINDA to improve its usefulness for data storage and retrieval is one example of the need for consultations between data centers.

V. Relationships between the INDC, the IAEA and the Nuclear Data Unit

With the establishment of the INDC and its Terms of Reference, and with the Nuclear Data Unit moving into the position of a full-time (fully occupied) data center, it becomes essential to examine the interactions between the INDC and the Nuclear Data Unit and, indeed, with the IAEA in general. It is not the purpose here to analyze what this relationship should be in the long view; it is the purpose to bring together relevant points from the preceding discussions in order to understand possible options that exist.

As was the case when center-to-center relationships were considered, under the topic heading The International System of Data Exchange, it is most essential to clearly distinguish between long range and short range points of view when considering relationships with INDC. In this case, in which we consider relations between the data centers and INDC, we confine our remarks to the immediate future for which certain initial and boundary conditions are known and certain most plausible assumptions can be made. Thus at the present time there exist four separate data centers whose systems, philosophies, objectives, etc., are essentially independent. We assume that for some considerable time to come, this will be the prevailing situation. While argument can be made to unify the operations of these centers, the opinion of the Nuclear Data Unit is that such an objective at the present time is both unwise and impractical. It is unwise because the whole art of information storage, retrieval and exchange is in much too early a stage of development and depends upon too many elusive factors to warrant choice of "a system" or even for too strong coupling between individual centers whose separate systems deserve exploration. However, it is possible to say from experience that four data centers cannot have a functioning data exchange without some relatively frequent (at least yearly) consultation between representatives of the centers. The results of such consultations should necessarily be reported to some group or committee competent to assay the resulting recommendations; here we propose that the latter function can be performed by the INDC. ~~IAEA through the INDC.~~ It seems clear then that the problems and policies of the individual data centers are distinct from the problems and policies by which the individual centers interact with one another. Having proposed the exchange of basic neutron information as a means of accomplishing more basic measurements suitable to meet the needs of reactor physicists and designers through international collaboration, and having established a system of data centers by which this can be achieved, the INDC has now as one of its most important tasks, to encourage and follow further developments of this undertaking.

It might appear in the context of the preceding discussions that the Nuclear Data Unit proposes a degree of independence of the INDC but this is not the intent. It is intended, however, to point out that data-center activities are separable from other activities and require reasonably predictable staffing requirements. If the INDC interacts with the IAEA through the Nuclear Data Unit, then other activities should be considered with respect to their specific staff requirements. Thus, for example, if the IAEA is to take an active position in regard to "nuclear standards", then any recommendations that the Nuclear Data Unit assume such a responsibility, should be in terms of additional staff for the purpose. It must be recognized that the new INDC Terms of Reference do <u>not</u> relate this committee and its activities to the Nuclear Data Unit of the IAEA but to the IAEA directly. In this enlarged relationship it is possible to visualize broadened activities of the IAEA in the field of basic nuclear measurements without necessarily increasing the size of the data compilation unit.

VI. SUMMARY:

In the preceding report, the Nuclear Data Unit has set forth an outline of its present activities, together with certain recommendations about the near future and certain speculations about the distant future. The paper contains much that will be controversial, but the position taken by the Nuclear Data Unit in all such matters is based upon considerable consultation and experience. The purpose of the paper nevertheless is to provide the INDC with a basis for discussion and for future considerations. It is perhaps fortunate that of the several topics covered, the following two positive statements can be made: a) the storage and retrieval system chosen for the IAEA is well adapted for the functions it has to perform according to all available information b) the staff and facilities requirements are fairly well determined with respect to the volume of data exchange.

A summary of the principal points contained in the report follows:

1. The present activities of the Nuclear Data Unit can be listed as a) neutron data compilation and exchange b) meetings c) standards.
2. The most important of the activities of the Nuclear Data Unit is data compilation and exchange.
3. The existence of a data center at IAEA is justified chiefly on the basis that it is one center in a <u>system</u> of data centers for achieving international data exchange for which four principal centers is about optimum.
4. The <u>system</u> of data centers for data exchange is an experiment whose <u>ultimate</u> success will be judged in terms of manifestations of its existence through producers and users of data.
5. It is premature to plan strong centralization either in terms of a "master-center" or in uniformity of centers.
6. The best result in the technical aspect of data storage and retrieval will result from <u>individual data centers</u> loosely coupled for exchange purposes with emphasis on steady growth of the system as a whole <u>simultaneous</u> with technical advance.
7. Close and systematic contact between the data centers is essential.
8. Short term goals by which success can be measured are essential. A sufficient if not necessary requirement of such a goal is that it be unattainable by conventional means (journal publication).
9. Such a goal is prompt availability of <u>all current results.</u>

10. A secondary short term measure of success is the volume of data exchanged (not necessarily volume of data stored).
11. To better achieve short term success (volume of exchange of data), a limited field of neutron data should be wisely selected; data on fissile nuclides is a favourable choice from the standpoint of 9. to 10.
12. A second strategic class of data is "evaluated data" (not evaluation of data) which could be treated in the same manner as experimental data without confusion.
13. It is possible to estimate in a reasonable way, staff and facilities requirements if the volume of data acquired and exchanged is known.
14. International data exchange cannot be a partial success - it completely succeeds or it fails.
15. Compilation activities can be readily extended to include facilities and standards.
16. If activities in the area of "nuclear standards" are to be a part of the activity of the Nuclear Data Unit, separate consideration should be given to its objectives and staff requirements.
17. The Evaluation of the 2200 m/sec fission constants is to be considered a standards activity.
18. There is need to extend "standards evaluation" beyond the evaluations of the 2200 m/sec fission constants and include several other "standards reactions" as well.
19. If meetings of the scope of seminars and conferences are to be a feature of the Nuclear Data Unit's activities, it is desirable for efficiency to make typical ones periodic and to reserve blocks of time for non-typical ones. Alternatively, it is almost necessary to specify a given meeting three years in advance.

ANNEX I.*

DATA STORAGE AND RETRIEVAL.

I. General Description of the System

The neutron data compilation activity at the IAEA has a dual purpose:

- to collect and disseminate neutron data within a geographical area that includes all countries outside North America, Western Europe, the USSR and Japan, and
- to participate with the other three major compilation centers, in the promotion of neutron data exchange on an international basis.

To carry out this dual objective, the IAEA Nuclear Data Unit has devised a processing system, composed of two computer systems coupled with a systematic bookkeeping and data handling procedure. This data processing system is based on the principle that numerical data and bibliographical information are stored on two separate and distinct physical files. The two programs designed to process the numerical and bibliographic data are called DASTAR and CINDU respectively.

A. DASTAR

The DASTAR (DAta STorage And Retrieval) program, designed to store and retrieve numerical data safely and efficiently, has the following general characteristics:

- Any data set originated from a given experiment is stored on tape in the order of its accession number, referred to at the Nuclear Data Unit as the DASTAR number.
- Data is stored in an unmerged state; that is, data from different experiments pertinent to a given isotope and reaction type are not merged.
- Each numerical table, representing a given experiment is preceded by a comment block which gives specific information pertinent to the experiment.
- Each numerical table, defined by its own DASTAR number (i.e., accession number), is unlimited in length.
- The format of stored numerical data is flexible, that is, as far as possible tabular data are stored in the format in which they were supplied to the Nuclear Data Unit.

---

* This supplements Annex I of Newsletter No. 1, 1967

B. CINDU

The CINDU (Catalogue of data stored at the IAEA Nuclear Data Unit) program is designed to store bibliographical data associated with the numerical data stored in DASTAR and to produce a catalogue of that data. It has the following general characteristics:

- All data sets contained in the DASTAR data file are referenced in the CINDU catalogue.
- The format of the CINDU catalogue is fashioned after the CINDA format.
- CINDU differs from CINDA in the following three aspects:
    - All references pertinent to a given experiment are blocked together.
    - All information is sorted on isotope, quantity, laboratory and year of experiment.
    - DASTAR accession number and secondary references appear in fixed fields.
    - Author names are entered in such a way so as to facilitate retrieval by author.
- Data which exist in other data storage systems can easily be linked to the CINDU catalogue.
- CINDU, published periodically, provides an up-to-date catalogue of data holdings at the IAEA Nuclear Data Unit.
- Retrieval of data from the DASTAR file is done primarily in the CINDU bibliographic part of the system; retrieval in CINDU can be performed on any item which is assigned to a fixed field (see Section II-D-2).

II. Data Processing

The processing of data at the Nuclear Data Unit consists of four specific functions:

- Requesting of data
- Acquisition of data
- Editing of data
- Dissemination of data

In principle, this system is based on the request and reply method, which is one of the mechanisms by which the neutron data library of the Nuclear Data Unit is designed to grow in volume. The primary method consists of a systematic collection by the Nuclear Data Unit of those data which originate within the service area of the Nuclear Data Unit. Another method, which in the initial stages of the Vienna Data Center operation has had a considerable effect on the data holdings of the Nuclear Data Unit, is the "Post Box" method of voluntary submission of data by the other 3 major centers, to promote international data exchange.

The processing of the data, however, is essentially the same, whether data is acquired by request, by collection or by voluntary submission, primary consideration is given to accuracy and efficiency, and to the prompt attention to requests.

A. Requesting of data.

Requests for data originate either within the NDU service area or from other data centers. Each request submitted to, and sent out by, the Nuclear Data Unit is assigned a number and entered in a request log, in which subsequent disposition and action, including all correspondence pertinent to the given request is recorded. Although presently in the form of a loose-leaf book, the request log will be computerized so as to facilitate the monitoring of requests.

The requests can be categorized as incoming and outgoing. When an incoming request is received, the CINDU catalogue is scanned to see if the data are on hand. If they are on hand, the data are retrieved from the DASTAR file and sent out in the format desired by the requestor. If the desired data set is not on hand, an outgoing request is originated and sent to either the appropriate data center, or to the author of the data if the place of origin is within the Nuclear Data Unit service area. In the event that the data requested have not been received within a 2-month period after the original request date, a follow-up request is re-issued. All request correspondence to individual authors, except when it is addressed to other centers, are accompanied by a Neutron Data Compilation Information Sheet.

B. Acquisition of data.

This part of the data processing includes all functions performed from the time of physical acquisition of data to when the data have been entered into the DASTAR system. When a data set is received by the Nuclear Data Unit, it is assigned to one of the data center physicists who sees to it that the data set is processed through the system in an efficient manner.

1. Initial Processing

When a data set is received, identifying information is entered into the DASTAR log, and an acquisition number is assigned to the set. For the purposes of data handling at the Nuclear Data Unit, a data set is defined as one consistent set, which originated from one experiment, is pertinent to a single isotope and reaction, and is continuous within a given energy region. In the event of an experiment which consisted of a number of individual runs, the data from each run are considered as separate data sets. The subsequent initial processing steps are the following:

(a) A copy of the original table is made, and the original is filed in the "originals" file. In the event that the information was supplied in the form of cards, they are duplicated, and the original cards are stored in a special card file. If the data were supplied on tape, the tape is listed and the list is filed in the "originals" file. In all cases, the original information is kept intact, and only copies of the original information are used in the data processing.

(b) A DASTAR folder is assigned to each set of data. This folder contains all pertinent material of a given data set, i.e. copy of original information, subsequent computer listings, etc. A Data Processing Form, attached to the folder, is designed to provide a continuous indication as to the processing status of the data set.

(c) A check with the request log is made to ascertain whether the given data set is on request. If it is on request, an appropriate note is attached to the DASTAR folder, in order to fill the request after the data has been processed through the system.

(d) When a data set is received in the form of a list of numbers only, without any accompanying descriptive information, a request is made for a copy of the primary reference from the IAEA library. This reference copy is then kept in the DASTAR folder together with the numerical data. Sometimes, when no information on a given set of data is available, it is necessary to hold processing until information is obtained from the author.

2. Documentation

Any information designed to identify a given set of data, to complete its bibliographic reference, and to annotate the data, is defined here as documentation. This step consists of the following functions:

- Composition of the DASTAR heading
- Making the appropriate CINDU entries
- Making the appropriate CINDA entries

(a) The DASTAR heading consists of two types of information: identifying and explanatory.

- Identifying information consists of:

(i) DASTAR number and version (see Editing of data, section C.2 below).

(ii) Type of experiment, isotope, energy range.

(iii) Author, place of origin and date of experiment.

(iv) How, when and under what circumstances the data was received by the Nuclear Data Unit.

(v) Identification of the variables appearing in each column of the data set.

(vi) Description of format, including the number of data lines, the number of variables per data line, and the format of each data line in FORTRAN notation.

- Explanatory information is designed to include all annotations, received from the author or extracted from literature, pertinent to the set of data. This includes descriptions of experimental apparatus, targets, detectors, and experimental methods, as well as information on error analysis and normalization procedures.
- Classification status of the data is also included in the DASTAR heading block; this may qualify the release of the data or indicate whether any particular set is considered by the author to be preliminary or final.

  The DASTAR heading information block is entered in alphanumeric form in a free-field format into a DASTAR-heading form which is subsequently punched into cards and combined on tape with the appropriate data set.

(b) The CINDU entries constitute the bibliographical information pertinent to a given set of data. Most of the information in CINDU, which appears in a fixed field can be used for retrieval purposes, this includes potentially all information, as well as author's name, with the exception of commentary in columns 44-79.

The CINDU entries are attempted to list the bibliographic references as completely as possible; this sometimes necessitates extensive literature searches which supplement information extracted from CINDA, and that received from the author.

(c) Necessary CINDA entries are made in parallel with the CINDU entries. When the information is already in CINDA, it is sometimes necessary to revise it, bring it up to date, or consolidate it. When the information is not in CINDA, a new entry is initiated and entered into the CINDA processing channel.

All DASTAR, CINDU and CINDA entries, after having been entered into specifically designed entry sheets, are double checked by another data center physicist. Only after it has been determined that all coded information is clearly understood and free of apparent errors, is it submitted for machine processing.

3. Machine Processing

Following all pre-processing and documentation steps, the DASTAR folder containing all pertinent information (data, reference copies, entry forms, etc.) is submitted to the machine processing. This stage of processing is performed with the aid of the IAEA computer facility.

(a) Preparation of information for input. Bibliographic and descriptive information is transcribed from the coded entry forms to punched cards, and verified.

(b) Data, although entered in the same format as it is received, necessarily needs some transformation in order to enter it into the DASTAR system:

- If received in the form of lists or tables, the data have to be punched on cards and verified.
- If received in Western format cards or tape, it may be necessary to transform the numerical format or separate merged data sets.
- If received in USSR format cards, which is the standard method of data transfer from the Obninsk Center to the Nuclear Data Unit, the data have to be translated to Western format cards.*

(c) After the necessary transcriptions and transformation of data have been completed, the DASTAR data sets, combined with their headings, are listed by the computer and returned, together with the DASTAR folder, to the responsible physicist for checking. This final checking procedure, before the updating of the DASTAR master file, consists of a check by the physicist assigned to the given data set, and of a secondary check by another physicist of the data center. This dual Check has proven beneficial in ascertaining the correctness of the information, as well as in the discovery of inadvertent errors.

(d) Following the final check, the DASTAR folder is returned to the computer processing section. Each DASTAR is then processed through the updating program which enters it into the DASTAR master tape in order of ascending DASTAR number. As a final step in the DASTAR machine processing, the computer retrieves the newly entered DASTAR and produces a hard

---

* Two computer programs have been written for this purpose:

TOPSI translates from USSR to Western format
ISPOT translates from Western to USSR format

copy list, with as many duplicate copies as necessary; the final DASTAR copies together with the DASTAR folder are then returned for final disposition to the physicist in charge, at which point a copy of the data list is sent to the author.

(e) The CINDU entries after having been punched on cards and verified, are processed through the CINDU updating periodically (approximately once a month) rather than consecutively as in the case of DASTAR. After a number of new CINDU entries have accumulated, they are merged into the CINDU master tape by isotope, quantity, laboratory, and year of experiment, in that order. A preliminary updated CINDU catalogue is then produced by the computer and distributed among the data center physicists for checking.

C. Checking and corrections procedures.

1. DASTAR checking procedures

As has been pointed out above, the processing of data through the DASTAR system is subjected to independent checks to ensure that the finally processed data and associated comments have a minimum number of errors. To summarize, this includes:

- An independent check of all coded entry forms by another physicist before submission for machine processing;
- A check of the preliminary processed information, in the final format, before updating of the master tape;
- A final check of every DASTAR data set, in the form in which it is stored at the IAEA Nuclear Data Unit, is performed by the author of the data. After the completion of the processing, a proof-copy of every DASTAR table is sent with a cover letter to the author.

2. DASTAR correction procedures.

Two methods are currently employed in handling corrections of DASTAR data sets which have been entered into the DASTAR library tape.

- If a DASTAR data set has been found by the author to contain mistakes, and a new version is supplied to the Nuclear Data Unit, the old DASTAR table is replaced by the new DASTAR table, given a new version number (see first line of DASTAR heading), and a note to that effect inserted into the DASTAR heading. The old version is discarded.
- If a DASTAR data set has been superseded by a new version, which originated as a result of re-evaluation, re-normalization, or other manipulation of the data, the new version is given a new DASTAR number, and the old version

is kept under the original DASTAR number. Both versions are cross-referenced, and an appropriate no-release flag is attached to the reference to the original version in CINDU. This method guarantees the retention of an old but possibly useful set of data, but prevents the free release of the old data unless specifically requested.

3. CINDU checking and updating procedure.

As was described in section 3 above, the CINDU master file is updated periodically after a sufficient number of new entries have accumulated. At each update, a new preliminary CINDU catalogue is produced and distributed among the data center physicists for checking and correction.

Each line in CINDU can be addressed by its line number, so that correction of CINDU is done simply by removing a given line of information, and replacing it by a corrected line. These interim updated versions of CINDU are primarily for the internal use of the data center; however, every few months, an updated CINDU catalogue is produced for external distribution.

D. Dissemination of data.

1. Initiation by request

All data stored at the IAEA Nuclear Data Unit are available to anyone without restriction, but the dissemination of these data is initiated by request only.

Every request received by the Nuclear Data Unit is filed chronologically, together with all information pertinent to its disposition. The request log (see section II.A) is continually monitored in order to ascertain that every request is answered as soon as the data are available. As was pointed out above, this procedure is presently performed manually; it is, however, anticipated to devise a computerized method which would couple the CINDU catalogue of data available with the current request log.

If a requested data set is not in the files of the Nuclear Data Unit, an outgoing request is sent to the author of the data, if the origin of the data is within the Nuclear Data Unit service area; if the origin is outside its service area, the data are requested from the appropriate data center. The subsequent acquisition and processing procedures are described in section II above.

If the requested data are on hand, they are retrieved from the DASTAR master file, and sent to the requestor.

2. Retrieval Procedures

The retrieval of data from the data storage system is a computerized procedure performed in two consecutive steps: the first step is a selective search of the CINDU master tape which yields the DASTAR numbers of the desired data sets together with all relevant bibliographical information; the second step is a straight-forward retrieval of the desired data sets, called by their DASTAR numbers, from the DASTAR master tape.

The retrieval search of the CINDU tape can be performed on any item, or a combination thereof, which are assigned to a fixed field in the CINDU format. At the present time, this capability includes all fixed field variables except for author's name, that is: Isotope (Z and A), reaction type (quantity), energy, laboratory, year of experiment, journal and date of publication.

The result of the CINDU retrieval search is a printed list of all bibliographic references pertinent to the data requested and a set of punched cards containing DASTAR numbers, devoid of duplications, and sorted in numerical order, ready for the subsequent retrieval from the DASTAR master file.

The desired DASTAR data sets are retrieved from the DASTAR master tape and written on a separate output tape which is then either sent to the requestor or used for producing the data in the desired output medium.

3. Output media

The requested data can be provided to requestors in a variety of formats:

a) Printed list: standard computer print-out of the full DASTAR table
b) Punched cards:
    (i) in IBM BCD format, with the alphanumeric comment block
    (ii) in USSR format, without the alphanumeric comment block
c) Magnetic tape: standard 7-track IBM in BCD format (800 or 556 BPI density)
d) Graphical plots: produced by a Calcomp digital plotter.