Nuclear Data Tagging in INIS

H.D. Lemmel
IAEA Nuclear Data Section
5.3. 1976

Contents:
1. Introduction
2. The proposals
3. How do the nuclear data centers obtain nuclear data?
4. Literature scanning by the nuclear data centers
5. Other user needs?
6. The basic differences between CIIDA and INIS
7. Past experiences with INIS retrievals
8. Principles of nuclear data indicators and tags
9. Additional requirements
10. Data tagging and flagging in an interdisciplinary frame

Annex: Proposals for nuclear data tagging in INIS

I. Subfield on data status
II. Subfield on data categories
Nuclear Data Tagging in INIS

H.D. Lemmel
IAEA Nuclear Data Section
5.3. 1976

Contents:
1. Introduction
2. The proposals
3. How do the nuclear data centers obtain nuclear data?
4. Literature scanning by the nuclear data centers
5. Other user needs?
6. The basic differences between CINDA and INIS
7. Past experiences with INIS retrievals
8. Principles of nuclear data indicators and tags
9. Additional requirements
10. Data tagging and flagging in an interdisciplinary frame

Annex: Proposals for nuclear data tagging in INIS
   I. Subfield on data status
   II. Subfield on data categories
1. Introduction

Specialized nuclear data centers and general nuclear science bibliographic information systems have part of their information sources in common; that is the nuclear data literature which they both cover. Since many years the relationship between data centers and bibliographic information systems has been rather loose.

The development of computer storage and retrieval systems, including the international organization of system input and information dissemination, seems now to have reached a stage where some well-defined coordination between INIS and the nuclear data centers appears to be desirable and also technically feasible.

A variety of nuclear data centers or continuing data compilation activities exist, each of them dealing with a broader or narrower specialized data scope such as isotopic constants, photonuclear data, neutron nuclear data, charged particle nuclear data and others. Some of these centers maintain, in addition to their numerical data files, a computerized data index and bibliographic file. A typical example is CINDA, which covers the field of microscopic neutron nuclear data.

For the production of CINDA and other specialized nuclear data reference files, NSA and INIS output have mainly been used for completeness cross-checks. More recently regular INIS retrievals are used by some of the nuclear data centers.

In view of the progressing development of INIS and its retrieval facilities, it seems now that the relations of INIS to nuclear data information, and of the Nuclear Data Centers to INIS services should be investigated in order to realize where improvements could and should be made.

Within such investigations the aspects of practicability and cost-effectiveness must be given due consideration.

2. The proposals

The Consultants' Meeting on Charged Particle Nuclear Data Compilation, held in Vienna 8-12 September 1975 [INDC(NDS)-69], recommended "to investigate the feasibility of deriving input for a nuclear data bibliographic file from a more general bibliographic system, especially from INIS".

Based on this recommendation the IAEA received from ZAED (telex of 24 Oct. 1975) a proposal to study the following items in the INIS system:

- a literary indicator for data and data compilations,
- the assigning of already available descriptors with regard to data,
- a special tag which would enable comprehensive and specific identification of data.

This recommendation and this proposal, which are not particularly related to charged-particle nuclear data but could have emerged from any other meeting on nuclear data, require some background information on the activities of the nuclear data centers. In the following it will be discussed how nuclear
data centers in general, and in particular neutron nuclear data centers, can best benefit from such proposals.

3. How do the nuclear data centers obtain numerical data?

It is essential to realize that for the nuclear data centers the main source for numerical nuclear data is not the printed materials referenced in INIS but the authors themselves. Many nuclear data, in particular large data sets, are no longer published in printed form and are exclusively available from the numerical data files held by the nuclear data centers. The pertinent literature which contains the detailed experimental description, error analysis, and other essential information needed for assessment and evaluation of the data, is however as important as the numerical data themselves. The data users usually require to have the numerical data plus the pertinent literature.

The nuclear data centers are supposed to have the numerical data in their files already before they appear in the literature. When an author prepares a lab-report or submits a paper to a journal, this is the time when the nuclear data center is supposed to compile the relevant numerical data. The center is supposed to be in contact with the author long before the experiment is finished; progress reports and preliminary results are essential. The center may even have stimulated the nuclear data measurement through the program compilation of nuclear data measurement requests. Close contacts to the scientific community of nuclear data producers and users are the basis for the nuclear data center activities.

The ideal practice of data compilation as outlined is performed only for data of primary importance, and some of the numerical nuclear data are extracted from the printed literature. But even in this case the center gets in contact with the authors to obtain additional unpublished relevant information and to verify whether the numbers published are still up-to-date. Occasionally the data center reviews the data and may advise the author to revise the data with respect to standard reference values used or corrections applied.

4. Literature scanning by the nuclear data centers

Data centers scan the literature for the following purposes:
- to extract numerical nuclear data if these have not yet been obtained directly from the author;
- to extract, for data sets already compiled, important supplementary information on experiment and analysis by which the data were obtained;
- to compile for each data set, a complete bibliography including progress-reports, lab-reports, conference papers, theses, journal articles, etc.;
- to obtain knowledge on experiments being started (progress-reports!) in order to contact the author.

Literature scanning is a time-consuming job which should be avoided if it can be replaced by INIS services. However, for the more important nuclear physics periodicals and report series which are dealing with nuclear data in almost every issue, it will be more efficient if the centers continue regular literature scanning. An access through INIS would be slower than the direct access, and the
papers themselves must anyway be studied. It is therefore not sufficient to simply transform an item retrieved from INIS into CINDA input format without studying the paper itself.

There is however a large number of periodicals and report series which contain nuclear data information only occasionally. For this literature of secondary relevance to nuclear data, literature scanning by the nuclear data centers is inefficient.

Conclusion:

The nuclear data centers would much benefit when they could stop scanning literature of secondary relevance with respect to nuclear data, and when they could rely upon this type of literature on regular INIS retrievals. How INIS could provide such retrievals to the nuclear data centers, will be the topic to be discussed in this paper.

5. Other user needs?

The conclusion just mentioned shows that the most important INIS user with respect to nuclear data reference information could be the nuclear data centers themselves. If any other customer somewhere in an institute or in industry is interested in nuclear data, it seems obvious that he would contact the data centers rather than INIS. For this purpose it may be helpful if the services of the existing data centers were advertised in the INIS Atomindex.

There may be some other interests by external customers related to nuclear data, for example questions about nuclear data measurement facilities or methods. If such or similar questions arise, INIS may quite well be prepared to give the answer using the presently existing descriptors. However, the actual questions submitted by customers should be analyzed before details can usefully be discussed.

6. The basic differences between CINDA and INIS

Data centers usually maintain a data index file, which is a prerequisite for their data compilation, review and dissemination activities. Frequently, an extract from this file is published in order to inform the customers of the contents of the data files available. Such publications vary in contents and format. Examples are newsletters (e.g. by NNDC and MDCC) listing new literature of interest, and indexing data sets recently received; the journal "Nuclear Data Sheets" which publishes in some issues evaluated nuclear structure data, in other issues "Recent References" relevant to such data; the data index CINDU of the IAEA Nuclear Data Section documenting in its last issue some thirty specialized numerical nuclear data libraries presently available; and last not least the neutron nuclear data handbook CINDA which will be used here as an example to be compared with INIS.

a) Data index versus bibliography

Although CINDA contains much bibliographic information, it is basically not a bibliography but a data index.

In a bibliography like INIS the entry unit refers to a document or paper. Each entry contains bibliographic information plus information about the contents in the form of an abstract and in indexed form to be used for computer retrievals.
In CINDA the entry unit is a numerical data set plus information about the data set. For each data set the following items are given: nuclide, reaction, energy-range, lab of origin, a list of all pertinent bibliographic references plus a comment what kind of information about the data set each reference contains, availability of the numerical data in a computer library, and some other items. Many CINDA entries also include an evaluation of references in the form of an indication or warning if a publication or a published data set is "preliminary" or "superseded".

b) Degree of specialization

CINDA is a highly specialized index to a data category which is called "neutron nuclear data". For INIS "neutron nuclear data" is only one out of dozens of other data categories. It seems to be unrealistic to assume that INIS can classify all these data categories to the same depth as CINDA does for its data category.

To give an example: In the numerical neutron data library EXFOR we have more than 75 different data-types describing different parameters of the process "neutron induced nuclear fission". Those data-types range from the basic fission cross-section as a function of the incident neutron energy, to highly specialized data like Legendre coefficients, in one of several possible representations, for the angular distribution of fission fragments, as a function of several variables.

In the data index CINDA, these 75 different data types are grouped in a set of 13 different "index-quantities" giving a less detailed classification of the neutron-induced fission data. This less detailed classification of data in the index has been organized according to the many-years experiences with the retrieval needs of the data users community.

The detailed level of classification of nuclear data which should be adopted in INIS will have to be analyzed according to the limits of practicability imposed on the precision cut-off within the wide scope of INIS, and according to the foreseeable retrieval needs of the nuclear data centers as the main users of INIS in the field of nuclear data.

c) Compact handbook or computer retrievals

The users of CINDA continuously stress the particular value of CINDA as a compact cumulative and conveniently structured handy reference book. This could not conveniently be replaced by a service of computer retrievals, although sometimes computer retrievals from the CINDA file are useful supplements to the printed books.

7. Past experiences with INIS retrievals

Since quite some time the IAEA Nuclear Data Section regularly scans certain subject categories, the report-number index and the conference index of the INIS Atomindex, and finds therein a lot of useful information within a greater lot of noise. This scanning is time-consuming but nevertheless brings out a significant amount of nuclear data information references. This way, the INIS Atomindex has become a useful tool for the nuclear data center activity, although the existing subject categories are not ideal for this purpose.

In addition, the Nuclear Data Section has a standing request for retrieving references on neutron nuclear data. The resulting retrievals are useful but not sufficiently reliable. The elementary retrieval needs of the Nuclear Data Section,
namely to retrieve all references dealing with the "determination of microscopic neutron nuclear data" can, at present, not be formulated adequately enough so as to avoid misses and noise. The term "microscopic neutron nuclear data" must be composed out of more specialized descriptors, and the descriptors available are apparently not comprehensive in the nuclear data field.

Also, the logical linkage between existing descriptors is often insufficient. A reference having, among others, the descriptors "neutrons", "fission products", "data", will often not include any "neutron-induced nuclear data on fission products" but, perhaps, some technical information on the use of neutrons and fission products plus some other kinds of data. This example illustrates the limitations in the use of the INIS descriptors for retrievals. The additional limitation is the naturally limited reliability of the INIS indexers entering the appropriate descriptors. The ZARD proposal seems to be suitable to overcome both of these limitations.

8. Principles of nuclear data indicators and tags

The ZARD proposal as quoted above includes two new elements to be introduced in INIS: an indicator flagging the presence of numerical data in the reference considered, and a tag for comprehensive and specific identification of data, both for the purpose of facilitating computer retrievals on INIS entries referencing nuclear data information.

According to the design of the INIS system, an indicator flagging the presence of numerical data can only carry the information "yes" or "no". Any more detailed information on the numerical data can be entered in a field which would be identified by a "numerical data tag" or, more specifically, a "nuclear data tag", (which would have the form of a specific 3-digit number.)

The desirable contents of this field identified by the nuclear data tag may be illustrated by the following considerations defining the retrieval needs of nuclear data centers. At least two subfields will be required: one subfield containing information about the data status, another subfield specifying the data category. Only the contents of these subfields will be considered, and not the way how these contents will be entered by means of cryptic or mnemonic codes or longer keywords. This question is of secondary importance.

8.a. The subfield on data status

The yes-no indicator for the presence of nuclear data does not meet the retrieval needs of the nuclear data centers. Some more information is required.

For the nuclear data centers it is essential to distinguish
- whether the numerical data given in the paper are the result of the work described, or
- whether the numerical data given in the paper are quoted from another source.

Furthermore, as explained above, the nuclear data centers are
- not only interested in retrieving references which give numerical results of nuclear data,
- but are as much interested in retrieving references dealing with the determination of nuclear data even without quoting results.

It should be noted that the result of a nuclear data measurement may consist of a single value which may be hidden in the text!

Thus, the type of information to be included in the first subfield would be:
1. determination of nuclear data without quoting results
2. determination of nuclear data with quoting results
3. nuclear data quoted from another source.
The comprehensive and specific identification of the different nuclear data categories seems to be the crucial question to be discussed. The question to solve is to invent a comprehensive classification scheme for nuclear data, detailed enough to meet the retrieval needs of the nuclear data centers, but not too detailed in order not to overburden the system and the indexers.

A basic classification scheme for nuclear data is the following:

A. "Intrinsic" nuclear data
   1. isotopic constants
   2. spontaneous decay data
   3. nuclear structure data

B. Nuclear reaction data
   1. photon nuclear data
   2. neutron nuclear data
   3. charged particles nuclear data.

This scheme corresponds roughly to the data scopes of existing nuclear data centers.

It should be noted that already with this basic classification scheme, a given bibliographic reference will often fall under two or more categories. Therefore, it must be possible that an INIS entry can include two or more strings of data-status/data-category subfields which can be retrieved separately.

Example: a paper describes the measurement of a neutron cross-section, using a half-life value from another source as reference value. This would require two strings of information:

<table>
<thead>
<tr>
<th>data status</th>
<th>data category</th>
</tr>
</thead>
<tbody>
<tr>
<td>determination with quoting results</td>
<td>neutron nuclear data</td>
</tr>
<tr>
<td>quotation from another source</td>
<td>nuclear decay data</td>
</tr>
</tbody>
</table>

This paper should not be retrieved, when retrieving on "determination with quoting results" of "nuclear decay data". This requirement is obvious and easy for this simple example, but is important to realize when the classification of flags and tags becomes more detailed and structured. If the nuclear data tag would have the depth of detail of a CINDA entry, one would have to be prepared to enter in INIS anything between one and more than one hundred such strings for a single reference.

3.c. If one includes in this consideration the INIS descriptors, the problem becomes even more complex, and one may have to enter, for each bibliographic reference, numerous strings of data-status/data-category/descriptors. For example, the INIS descriptor "activation detectors" may refer to the one kind of nuclear data mentioned in the paper but not to the other kind. Furthermore, nuclear data may be measured with activation detectors, or quoted for the use with activation detectors.

Such logical links between the tagged nuclear data field and the INIS descriptors may not be feasible, and I therefore leave the descriptors out of my considerations. However, several types of nuclear data are reflected by "intrinsic" INIS descriptors, which must be reviewed after introducing the new field on nuclear data information.
9. Additional requirements

For the compilation work of the nuclear data centers it would be essential to have INIS retrievals on nuclear data references also specified by geographical criteria. It should be possible to retrieve all nuclear data relevant information

- where the work was done
  - in a given laboratory, or
  - in a given country, or
  - in a given list of countries, or
  - excluding a given list of countries;

- where the publication originates from
  - a given country, or
  - a given list of countries, or
  - excluding a given list of countries, or
  - translation from another language.

10. Data tagging and flagging in an interdisciplinary frame

A number of nuclear data categories are closely related to other scientific disciplines such as solid state physics, chemistry, thermodynamics. It will therefore be useful to coordinate the tagging and flagging of nuclear data with similar efforts dealing with the tagging of other kinds of data. Several such efforts seem to exist.

IUPAC has issued a preliminary list (unpublished) of chemical data categories to be tagged in primary and secondary literature. This example illustrates the overlap of data categories in different disciplines. Among nearly 100 different chemical data categories, this list includes seven categories of nuclear data. These are, in alphabetical order:

- Cross-section data (all types)
- Isotopic data (nucleic masses, abundances, etc.)
- Neutron diffraction data
- Nuclear reactions data (cross-sections, resonance energies, etc.)
- Radiation energy (alpha, beta, gamma, etc.)
- Resonance parameters (energies, widths, integrals, potentials)
- Radioactivity data (half-life, abundance, energy)

Obviously, these data categories do not seem to be well defined since some of them are overlapping. They may well be improved if IUPAC would take over, for the chemical nuclear data field, some of the nuclear data tags to be adopted in INIS.

On the other hand, it will be useful for INIS to study the data tag interests of such user groups as represented by IUPAC. Furthermore, it seems to be advisable that INIS adopts existing data tagging schemes from other disciplines as far as such literature is abstracted in INIS.
In this annex, draft proposals for retrievable items to be entered in INIS in a field identified with the nuclear data tag, are presented. Since these are based on the experiences obtained in the IAEA Nuclear Data Section, which does not deal with all kinds of nuclear data, it will be advisable to circulate these proposals to other specialized nuclear data centers.

I. Subfield on data status

Two subclassifications of this subfield will be useful to nuclear data centers. These are

**Subfield I.a. determination or use of data**

1. determination of nuclear data by experiment
2. determination of nuclear data by theory
3. determination of nuclear data by evaluation

(Note: "data evaluation" means deducing recommended best data from all available experimental and/or theoretical information; the "evaluation" of final results from the "raw data" of an experiment should not be included.)

4. compilation of nuclear data from many sources
5. critical review of nuclear data (discussion of discrepancies)
6. nuclear data quoted from another source and used for some purpose.

**Subfield I.b. presentation of data**

a. work in progress but no results yet
b. work finished but results not quoted
c. results presented in graphical form
d. results presented in numerical form
II. Subfield on data categories

The complexity of nuclear data classification shall be illustrated by the following list of items required for a detailed specification of nuclear reaction data, not including "intrinsic" nuclear data,

1. the target, that is either: element number and mass of the target nucleus (plus occasionally an indicator for a metastable state) or, in the case of thermal neutron scattering: a chemical compound
2. the incident particle (the term "particle" shall include photons)
3. the energy or energy spectrum of the incident particle
4. the reaction, either specified as "absorption", "fission", etc. or specified by the particles produced
5. the residual nucleus and its state
6. the reaction parameter(s) considered (angles, energies etc. of the particles produced)
7. an indication defining to which of the produced particles the reaction parameter(s) refer
8. representation of the data (measured points, fitted parameters, etc.).

To deal with such detailed data specification is certainly the task of the nuclear data centers and not of INIS. But which of these items should be included under the INIS data tag?

As explained earlier, the most basic parameter needed in INIS retrievals for the nuclear data centers is item 2., the incident particle. But this is not sufficient. For example, the data scope of GINDA is

primarily: neutron-induced nuclear data (defined by item 2. above)
but also: fission gamma, fission neutrons, fission products from spontaneous fission or from photofission; photon neutrons; (defined by items 2., 4., and 7 above).

Another example: the data scope of the Charged-Particle Nuclear Data Group at Karlsruhe is defined by:

primarily: charged-particle induced nuclear data (defined by item 2. above)
but only: integral cross-sections (defined by item 6. above)

Since "integral cross-sections" are only a very small but important fraction of all existing charged particle nuclear data, this specification is, for this purpose, essential to include in the INIS nuclear data tag. It would, however, not be very useful to include this specification for all kinds of nuclear data. Instead, other specifications will be needed for other kinds of nuclear data.

I want to demonstrate with this example that a very logical and systematic classification of nuclear data is certainly not the most practical one. Instead, one needs a more pragmatic scheme according to the actual or foreseeable retrieval needs of the primary customers represented, in the present case, by the nuclear data centers.

The aim of this proposal is to facilitate INIS retrievals covering, with minimum noise and misses, the full scope of a given nuclear data center activity, such as GINDA or EXFOR, assuming that the references retrieved will be further digested by the data center for input into its systems. It is not assumed that this digesting activity can be assisted significantly if the INIS data tag would contain more specific detail. However, many further details will certainly be retrievable from INIS descriptors such as: isomeric nuclei, polarization, gamma spectra, keV range, etc.

On the other hand it is realized that not all existing nuclear data center activities have been considered sufficiently and that the definition of their scopes is likely to require further subdivision of the one or other data category included in the following list. Therefore, this proposal should be submitted to those data centers for which such kind of INIS retrievals would be of interest, in order to obtain their comments.
List of retrievable items to be entered under an INIS nuclear data tag

Draft proposal

Several items will have to be split up further if this is required for defining INIS retrievals covering the scope of specialized nuclear data centers.

A. Intrinsic nuclear data
   1. isotopic constants, masses, abundances, etc.
   2. spontaneous decay data
      a. half-lives
      b. radiation energies, intensities, etc. of α, β, γ decays
      c. spontaneous fission data (including fission gammas, fission neutrons, fission fragments)
   3. nuclear structure data: nuclear levels and their properties
      (Note that these data may be deduced from reaction data induced by photons, neutrons or charged particles.)

B. Nuclear reaction data
   1. photo-nuclear data
      a. photo-neutron data
      b. photofission data (including fission gammas, fission neutrons, fission fragments)
      c. other photo-nuclear reaction data
   2. neutron nuclear data
      a. cross-sections as function of neutron energy
      b. yields, angular distributions and energy spectra of one or more of the following products: photons, neutrons, protons, deuterons, Ne-3 particles, alpha particles, but excluding nuclear fission.
         (Note: including thermal neutron scattering but excluding solid state physics data, neutron diffraction data, etc.)
      c. yields, angular distributions and energy spectra of fission gammas, fission neutrons, fission fragments
      d. resonance parameters, strength-functions, etc
      e. theoretical model parameters
      f. group data and scattering matrices
   3. charged-particle induced nuclear reaction data
      (Note: the term "charged particles" includes protons, deuterons, tritons, Ne-3 nuclei, alpha-particles)
      a. integral cross-sections (excitation functions)
         (Note: including integral cross-sections obtained from integration of differential data)
      b. thick-target yield data
      c. differential data
         (Note: often used for deriving nuclear structure data A.3.)

C. Data related to Nuclear Data
   1. elementary particles reaction data (high energy physics)
   2. nuclear reaction data involving other elementary particles than mentioned under B. above (for example meson production)
   3. macroscopic neutron cross-sections, neutron diffusion data, etc.
   4. solid state physics data from X-ray and neutron investigations; neutron diffraction data
   5. photon-electron interaction data in the atomic shell
   6. electron and positron interaction data in the atomic shell
   7. reaction data of atoms, molecules and ions
      a. for thermonuclear fusion problems
      b. other
In the given list of retrievable nuclear data items to be tagged, the retrieval criteria corresponding to the scope of CINDA would be:

B.2. neutron-nuclear data
   excluding B.2.f. group data
A.2.e. spontaneous fission data
B.1.a. photoneutron data
B.2.b. photofission data
but excluding data status 6: "data quoted for applications only."

The retrieval criteria corresponding to the scope of neutron nuclear data compilation in the EXFOR system, would be:

B.2. neutron nuclear data
   excluding B.2.c. theoretical model parameters
   excluding B.2.f. group data
A.2.c. spontaneous fission data
but including only data status 1: "determination of nuclear data by experiment."

For a system as general as INIS, this list should certainly be expanded by including all the other numerical data relevant to the INIS scope such as thermodynamic data, etc. It seems advisable to give priority to improving the classification of such data categories for which no specialized data centers exist and for which, therefore, the printed materials referenced in INIS are the primary numerical data source for data users.
References and abbreviations

The "Four Neutron Nuclear Data Centers" are:

NMCSC  USA National Neutron Cross-Section Center at the
        Brookhaven National Laboratory

NDCC  NEA Neutron Data Compilation Centre at Saclay, France

NDS  IAEA Nuclear Data Section, Vienna

CJD  USSR Centr po Jadernym Dannym, at the Fiziko-Energeticheskij Institut,
      Obninsk

CINDA 75, an index to the literature on microscopic neutron data. Published on
       behalf of the Four Neutron Nuclear Data Centers, IAEA Vienna 1975,
       a cumulative issue in two volumes, and a supplement.

CINDU-11, a catalogue of numerical nuclear data available from the IAEA Nuclear
        Data Section. IAEA Vienna, to be issued in March 1976.

INDC(NDS)-69, report on the Consultants' Meeting on Charged Particle Nuclear Data
        Compilation, Vienna, 8-12 September 1975. IAEA Vienna,
        to be issued in March 1976.

EXFOR, a computerized system for the storage and exchange of numerical neutron
        nuclear data and associated information. The system is jointly
        operated by the Four Centers named above. Internally documented;
        compare EXFOR Annual, and CINDU-11.

WHENDA 75, world request list for nuclear data measurements for fission reactor
        programmes, fusion reactor programmes, nuclear safeguards programmes.
        Published on behalf of the Four Centers named above,
        INDC(SEC)-46, IAEA Vienna 1975.