Title
COMPUTERIZED PRODUCTION OF A DATA COMPILATION - A TABLE OF ISOTOPES

Permalink
https://escholarship.org/uc/item/4mw4v2z5

Author
Lederer, C. Michael.

Publication Date
1973-10-01
COMPUTERIZED PRODUCTION OF A DATA COMPILATION - A TABLE OF ISOTOPES

C. Michael Lederer

October 1973

Prepared for the U.S. Atomic Energy Commission under Contract W-7405-ENG-48

TWO-WEEK LOAN COPY

This is a Library Circulating Copy which may be borrowed for two weeks. For a personal retention copy, call Tech. Info. Division, Ext. 6782
DISCLAIMER

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.
I was pleased to accept the invitation to speak to you today, especially in view of the fact that one of the sponsors of this meeting, Technical Information Division, has provided us with invaluable help in our compilation work over a period of many years. I confess from the outset that my experience in the information processing field is both limited and specialized. Nonetheless, I believe that many of the problems we face in the course of producing a nuclear data compilation occur also in other areas of information and data processing. It is my hope that some of the solutions we are developing will find other applications. In particular, I shall describe a new input-editing system for complex, technical text. Before doing so, I should like to discuss the contents of the Table of Isotopes as it relates to the need for and the requirements placed on computer automation.

The most recent (6th) edition of the Table [1] was published in 1967; a page from it is shown in fig. 1. One notes that the information (on radioactive decay and nuclear level properties) is presented in a complex format, involving both text and graphics, and a variety of character types - Roman and Greek letters, subscripts and superscripts, and a variety of special symbols and type fonts. Because this "language", if one may call it that, is easily understood by scientists who use the compilation, it is highly desirable to retain these features of the presentation in the automation process. This of course implies a degree of complexity in the output techniques used to reproduce the compilation. What is less evident, is that it is also desirable to retain some of this complexity at the input stage. This should be done because the data compilers also speak in the scientific "language", and will produce a more accurate, error-free compilation if they can write the data in its natural symbols.

The quantitative aspects of the compilation work also exert an important influence on its automation. There are approximately several million characters of print in the Table that comprise several hundred-thousand pieces of numeric data on 1500 isotopes. There are several hundred line drawings of the type just shown that also include extensive numeric data, and ten-thousand reference citations. These magnitudes are not large enough to present serious problems or costs in computer storage or handling. However, they present formidable
problems in the abstracting, copying, input, and proofreading of the data. A major objective of the automation is to minimize the impact of these problems on the data compiler, so that he may concentrate his efforts on the selection and evaluation of new data from the several thousand relevant papers that are published annually; that is, on the essential intellectual work that must be done by a scientist.

There are a number of clearly defined tasks that we wish to delegate to the computer:

1) It must accept new data in a format easily remembered and used by the scientist-compiler, with a minimum of recopying.

2) It must store the data and permit easy retrieval, editing, and addition of new data.

3) It should check the data insofar as possible for correct syntax, for consistency with the laws, conventions, and probabilities of physical science, and for sensible correlations.

4) It should produce level-scheme drawings directly from physical input data, plus a bare minimum of graphic information that varies from one drawing to the next.

5) It must select and sort the data, and print (or plot) them in a form suitable for direct reproduction by photo-offset.

Automation of entry, storage, retrieval, editing, sorting, and printing has been done for a number of other types of data compilations - airline schedules, phone books, and even Playboy Magazine (with exception of the graphics). The difference between automation of the Table of Isotopes and of a phone book lies in items (3) and (4) above, and in the complexity of the "language" used.

It is the complex language that motivated us to develop a new input-editing system, a description of which I now turn to. The system is an outgrowth of earlier developmental work by L. P. Meissner and M. L. Clinnick at this laboratory, motivated in part by the Table of Isotopes project [2,3]. Figure 2 shows the input console in use. The keyboard is an expanded typewriter keyboard, and the display is a storage oscilloscope. The console is connected to a small computer, which can drive multiple terminals. The present system was developed by E. Romascan, W. Greiman, and myself.

Figure 3 is a schematic drawing of the keyboard layout. The character set includes Roman and Greek alphabets in upper and lower case, numbers, and a variety of special symbols. Fonts include "normal" and combinations of superscript or subscript, italics, and boldface. The following are some additional features of the system:

1) It permits easy editing, both when entering data and later, after the data have been entered and proofread. A moving cursor can be positioned anywhere in the text and insertions or deletions performed at the position of the cursor. (Insertions appear at the bottom of the screen, and are replaced in the proper part of the text by striking the redisplay key.)
2) Compound symbols, such as $e_\gamma$, can be entered with a single key. In fact, the typist can create such a compound symbol and assign it to a particular key while typing. Another reason for redefinition of keys is that it permits frequently-used "fonted" characters - a Greek letter or a subscript for example - to be typed without requiring the use of a font key also.

3) "Template", or prompting routines can be used to insure the inclusion of required data identifiers. The entry of data for the Table of Isotopes makes use of one such routine; the upper portion of fig. 4 shows an example of a completed template. Such a template is not required when entering running text, but if one wants to enter many short items of data into a structured data file, it is a useful device for identification of the data in a way that will insure proper flagging for later retrieval.

4) A true columnar format can be used to enter data directly from published tables. (In this format, positioning of the data in columns and rows is controlled in a simple fashion with the TAB and LINE FEED keys.) An example of such a table is shown in the lower portion of fig. 4. In our use of columnar format for the Table of Isotopes, the number of columns and the column headings (identifiers for the numeric quantities that appear in each column) are defined by the FORMAT identifier that is entered as part of the preceding template. Because of this feature, the data compiler may give large tables, taken directly from the journals, to the typist, without the need for recopying. Entries in columnar format, including the TABs and LINE FEEDs themselves, can be edited in the same manner as non-columnar input.

5) The small computer that drives the consoles is linked directly to the large computers of the laboratory's central computer facility. This is not an essential feature of the system. (In fact, the original prototype of the input-editor was built as a stand-alone, self-contained system.) However, at an installation that possesses a large computer, there are a number of special advantages derived from its use:

   a) Data storage facilities of the large computer, both temporary (magnetic disk) and permanent (magnetic tapes, data cells, and photo-digital chips) are available for use. Besides placing the entire data bank at the call of the console, these devices substitute for the most costly components of a stand-alone system.

   b) The user of each console has a separate program in the large computer. As a result many of the operational features - the absence or presence of a template routine and the contents of the template, the definition of tabs and/or columns, and special functions such as title centering and formula layout - can be different for each user.
c) The far greater computational capability of the large computer can be used to accomplish most of the difficult display formatting operations rapidly. More important, these operations can be programmed in FORTRAN instead of machine language, which greatly simplifies the software development problems.

d) The console can be used as a general-purpose on-line terminal, simply by pressing a mode-change key. In this mode, the console functions exactly as an on-line teletype terminal.

I should hasten to add that there are also disadvantages of dependence on a large computer system. They add operating costs. More important, such systems are inherently unreliable. In event of a failure, they are magnificently programmed to save all job input information, so that jobs that were in the system at the time of the failure can be rerun. This is of little use to a typist, who wants to save the data rather than to start typing all over again. However, we have found software solutions to these problems; I am happy to be able to report that, following some unpleasant early experiences, we are now able to recover data-in-process at the time of a failure with almost 100% reliability. In fact, the recovery is more reliable than the recovery of primary accounting information in the large computer, so that one unexpected compensation for the annoyance of machine failures has been a small economic benefit.

There are plans for several other applications of this input-editing system in the laboratory. The Particle Data Group in Physics Division, which produces a compilation of elementary particle data, already has a terminal, and plans to use it in the near future. Technical Information Division is involved in a pilot project to produce journal articles on yet another console. When the paper has been accepted and edited by the journal editors, they will enter the corrections and provide the final version in the form of a photographic print or a magnetic tape suitable for input to a photo-composer. The objective is to eliminate entirely the cost of composition and the need for the author to perform a second proofreading.

I should like to mention briefly several other aspects of computer application to the Table of Isotopes that I alluded to earlier. Work is in progress on sorting of textual data and conversion from the identifiers-plus-data input format; illustrated in fig. 4, to a final output format, such as is illustrated in fig. 5.

Work on the processing of graphic data has been nearly completed. Figure 6 shows a level-scheme plot produced by the computer. It should be obvious that the ability to generate such graphics with little or no graphic input information is dependent on a very stylized format, and that the very complex programs that make such plots must be designed for a particular kind of data and plot-style. There are, however, some generally useful features of such programs, such as the ability to draw lines around labels, and to adjust line positions to accommodate labels. Even without considering other applications, the one man-year of programming required to produce such plots will result in a saving of 4 man-years of drafting on each future edition of the Table. More important, the latest data can be plotted in a very short time; otherwise, we would be faced with the unpleasant alternatives of publishing data that is unacceptably out of date, or of finding enough draftsmen to compress the 4 man-year drafting job (as well as the proofreading job) into an acceptable length of time.
The use of the computer to check for errors of syntax and violations of physical laws, conventions, or probabilities likewise requires extensive programming work, and is also specialized to the particular kind of data with which one is working. However, the computer is an ideal tool to do such perfunctory and repetitive checking, and the use of it in this manner for other applications is easily visualized. For example, is the departure time keyed into an airline schedule a proper number, with a value between 0:00 and 12:00, and are its units AM or PM, rather than BM, aM, or #%? Is the arrival time, when corrected for any time-zone difference, later than the departure time by the correct amount for the particular distance and aircraft? Do any two phone subscribers have the same number listed (incorrectly) in the phone book? I do not know the extent to which such checking is done in other applications; I suspect that it is in fact done, but not to the full extent possible. A particularly important application of checking is the analysis of highly correlated data, such as appears on the level-scheme diagrams, and in the departure vs. arrival time example I have just mentioned.

In conclusion, I would like to express a few retrospective thoughts on the work we have done to computerize the Table of Isotopes. The goals and the basic philosophy with which we approached the problem were

1. To define clearly what the computer is capable of doing, as distinct from intellectual work that must be done by the compiler;
2. To apply the computer wherever possible, in such a way as to eliminate unnecessary, routine work and thus provide the compiler with a maximum amount of time to do what the computer cannot do;
3. To make the computer aid the compiler in checking and proofreading by detecting all errors of form or content that can be machine-detected, thus reducing errors in the final compilation.

These goals have proven sound and attainable in practice, and I believe they are good principles with which to approach any information-handling problem. The more subtle subject is the balance between the extent to which the full capabilities of the computer are implemented and the cost of doing so. Experience has shown that this balance requires an evaluation of the benefits of each step in the process vs. the difficulty of implementing that step. Mistakes are made, so that it is important to re-evaluate the balance as automation proceeds; otherwise the mistakes become costly and time-consuming. When we have erred, it has been almost always on the side of too much automation. However, it is one of the nice things about computers that, when used intelligently, or even when overused cleverly, they often open up the possibility of accomplishing new goals; I have been pleasantly surprised in this way more than once. It seems very likely that the real benefits of computers to our data compilation, and to information handling in general, will be measured in terms of better information in more useful forms, not in terms of cost-effectiveness.
FOOTNOTES AND REFERENCES

* Work performed under the auspices of the U. S. Atomic Energy Commission and the U. S. National Bureau of Standards.


FIGURE CAPTIONS

Fig. 1. A page from the sixth edition of the *Table of Isotopes*. Production was entirely by hand. The line drawings were done by draftsmen; text and labels on the drawings were typed on an IBM typewriter.

Fig. 2. The console of the input-editing system.

Fig. 3. Schematic layout of the keyboard. Special function keys in the top two rows are for

a) position control ("Tab", "Line feed", "New Page")

b) display control ("Last Page", "Next Page", "Go to" [page number] "Send", and "Redisplay")

c) editing (Cursor keys, "Delete left", "Reset cursor", "Clear data")

d) retrieval of a specified item ("Go to" [item number X] "Send") (X = f for identifier flags, d for data)

e) template control ("Advance Template")

f) redefinition of a key ("Create symbol" [keys] "on key" [key])

g) column number identification ("Define column")

h) fonts ("Subscript", "Superscript", "Italic", "Bold", "Greek"). Foot pedals can also be used for subscript or superscript.

Fig. 4. Sample output (proofcopy) of part of a data item typed on the input-editing system. The upper portion contains the data identifiers, typed in response to the questions "ISOTOPE":, "DATA CATEGORY":, etc. (Identifiers labelled with superscript numbers in brackets were defined during entry of a previous item.) The data (lower part of figure) is in columnar format, with the column headings given in the "FORMAT" identifier. _, T, and L are symbols for blank, tab, and line feed. These symbols are printed to facilitate editing.

Fig. 5. Example of how the final output format for textual data will appear. Such a block of text is derived from a number of different input data items, like the one illustrated in fig. 4.

Fig. 6. A computer-drawn level-scheme diagram. This format, produced from physical input data, is used for both proof and final copy.
ITEM: 630  
TAPE: 30028  
ISOTOPE: $^{115}$Ag-$^{47}(20_{\pm}m)$\cite{619}  
DATA CAT.: $\gamma$(keV)  
REFERENCE: NP.A143.289//70Hn01  
METHOD: Ge(Li)$_{scint}$-Ge(Li)$_{y}$$_{coinc}$-Ge(Li)$_{scint}$-Ge(Li)$_{y}$$_{coinc}$  
FORMAT: $E_{\gamma}\pm\Delta E_{\gamma},t_{\gamma}\pm\Delta t_{\gamma}$  
A OR C: A\cite{619}  
COMPILER: EB\cite{619}  
DATE: 11/20/72\cite{619}  
TYPIST: AMN\cite{619}  

<table>
<thead>
<tr>
<th></th>
<th>$E_{\gamma}$ (keV)</th>
<th>$t_{\gamma}$ (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>131.4$\pm0.2$</td>
<td>14.4$\pm0.6$</td>
</tr>
<tr>
<td>2</td>
<td>213.6$\pm0.2$</td>
<td>24.8$\pm0.6$</td>
</tr>
<tr>
<td>3</td>
<td>229.7$\pm0.2$</td>
<td>100</td>
</tr>
<tr>
<td>4</td>
<td>237.1$\pm0.4$</td>
<td>2.2$\pm0.2$</td>
</tr>
<tr>
<td>5</td>
<td>243.6$\pm0.4$</td>
<td>1.8$\pm0.3$</td>
</tr>
<tr>
<td>6</td>
<td>277.4$\pm0.6$</td>
<td>0.4$\pm0.2$</td>
</tr>
<tr>
<td>7</td>
<td>303.3$\pm0.2$</td>
<td>3.4$\pm0.2$</td>
</tr>
<tr>
<td>8</td>
<td>326.6$\pm0.2$</td>
<td>10.6$\pm0.8$</td>
</tr>
<tr>
<td>9</td>
<td>360.9$\pm0.2$</td>
<td>3.3$\pm0.4$</td>
</tr>
<tr>
<td>10</td>
<td>372.6$\pm0.2$</td>
<td>10.2$\pm0.8$</td>
</tr>
<tr>
<td>11</td>
<td>389.3$\pm0.2$</td>
<td>1.9$\pm0.3$</td>
</tr>
<tr>
<td>12</td>
<td>417.3$\pm0.4$</td>
<td>1.7$\pm0.6$</td>
</tr>
<tr>
<td>13</td>
<td>473.2$\pm0.2$</td>
<td>18.7$\pm0.9$</td>
</tr>
<tr>
<td>14</td>
<td>507.7$\pm0.2$</td>
<td>7.1$\pm0.7$</td>
</tr>
<tr>
<td>15</td>
<td>539.6$\pm0.3$</td>
<td>0.7$\pm0.2$</td>
</tr>
<tr>
<td>16</td>
<td>548.3$\pm0.3$</td>
<td>1.1$\pm0.2$</td>
</tr>
<tr>
<td>17</td>
<td>585.0$\pm0.3$</td>
<td>0.8$\pm0.2$</td>
</tr>
<tr>
<td>18</td>
<td>649.9$\pm0.2$</td>
<td>13.8$\pm0.9$</td>
</tr>
<tr>
<td>19</td>
<td>699.2$\pm0.2$</td>
<td>10.1$\pm0.8$</td>
</tr>
<tr>
<td>20</td>
<td>719.0$\pm0.4$</td>
<td>0.4$\pm0.2$</td>
</tr>
<tr>
<td>21</td>
<td>750.0$\pm0.4$</td>
<td>0.5$\pm0.2$</td>
</tr>
</tbody>
</table>
\[ t_{1/2} = \begin{array}{ll}
8.085 \text{d} & (\text{Iiso 22 715}) \\
8.054 \text{d} & (\text{PMB 2 255}) \\
8.067 \text{d} & (\text{PMB 2 360}) \\
8.073 \text{d} & (\text{Iiso 19 475}) \\
8.070 \text{d} & (\text{NSEg 32 46}) \\
\end{array} \]

no modification of \( t_{1/2} \) by chemical environment (PR C3 1699)

others (RRou 13 485, RRL 2 41, CJC 31 120, PR 90 443, Nucl 11n3 14, PR 81 643, Nat 167 365)

\( \beta^- \) (PR 54 775)

Class

A. \( \text{Ident}: \text{chem} \) (PR 54 775); \( \text{chem}, \text{genet} \) (PR 57 363)

Prod.

fission

(\( \text{PR 56} \), \( \text{Nwis 27 529} \), \( \text{PR 54 775} \), \( \text{Nat 158 163} \), \( \text{PR 57 363} \), NNES 9 984, CJR 25B 371, Nat 161 520, NNES 9 1368)

\[ I = 7/2; \mu = 2.738 \; \text{e}, \quad q = -0.40 \; \text{e}; \quad \text{atomic beam} \] (PR 119 2022)

\( \beta^- \)

\( 0.08; 1 \) (0.6%), 0.605 (86%), 0.336 (13%), mag (Phca 17 637)

0.807 \( \text{d} \) (1.4%), 0.6059 \( \text{d} \) (84.7%), 0.339 \( \text{d} \) (13.3%), mag (PR 86 863)

0.818 \( \text{d} \) (0.7%), 0.608 \( \text{d} \) (83%), 0.335 \( \text{d} \) (9%), 0.250 \( \text{d} \) (8%), mag, \( \beta^- \) coinc (PR 86 212)

average \( \beta^- \) energy: 0.19, ion ch (PR 86 82)

others (ZP 179 62, Phil 43 648, Phil 43 221, Nat 170 915, PR 86 82, PR 83 860, PR 82 103, PR 81 482, Phca 17 658, PR 78 179, PR 76 94, PR 75 1270, PR 74 1879, PR 74 1640, PR 61 686)

\( \gamma \)

(see also \( ^{131}\text{mXe} \))

0.08014 \( \gamma \), 0.284307 \( \gamma \), 0.364467 \( \gamma \), cryst (PR 91 1027)

0.080165 \( \gamma \), \( \left( \gamma_{1} = 2.72 \text{d} \right) \), \( \text{ex}/\gamma = 1.33 \text{d} \), 0.1639 \( \gamma \), \( \left( \text{K/L} = 1.83 \text{d} \right) \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 6.1 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 1.2 \text{d} \), with \( ^{131}\text{mXe} \)\)

0.017723 \( \gamma \), \( \left( \text{K/L} = 0.82 \text{d} \right) \), \( \text{ex}/\gamma = 1.05 \text{d} \), \( \text{K/L} = 5.5 \text{d} \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 1.5 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 2.7 \text{d} \), \( \text{L}_{\text{III}}/\text{L}_{\text{IV}} = 1.3 \text{d} \)

0.2723 \( \gamma \), \( \left( \text{ex}/\gamma = 0.9 \text{d} \right) \), \( \text{K/L} = 5.5 \text{d} \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 2.9 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 3.4 \text{d} \)

0.2500 \( \gamma \), \( \left( \text{ex}/\gamma = 0.9 \text{d} \right) \), \( \text{K/L} = 5.1 \text{d} \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 4.5 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 5.0 \text{d} \)

0.3500 \( \gamma \), \( \left( \text{ex}/\gamma = 0.9 \text{d} \right) \), \( \text{K/L} = 5.4 \text{d} \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 3.4 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 1.7 \text{d} \)

0.3500 \( \gamma \), \( \left( \text{ex}/\gamma = 0.9 \text{d} \right) \), \( \text{K/L} = 5.4 \text{d} \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 3.4 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 1.7 \text{d} \)

0.4500 \( \gamma \), \( \left( \text{ex}/\gamma = 0.9 \text{d} \right) \), \( \text{K/L} = 5.4 \text{d} \), \( \text{L}_{\text{I}}/\text{L}_{\text{II}} = 3.4 \text{d} \), \( \text{L}_{\text{II}}/\text{L}_{\text{III}} = 1.7 \text{d} \)

mag conv (APHu 28 13)

others (RaAc 10 1, CR C264 944, ZP 179 62, NP 43 650, NP 40 566, NP 31 456, NP 24 318, Magy 2 n3, IzF 23 206, PR 101 746, ArkF 21, Phca 20 243, PR 90 849, CJP 30 715, PR 86 884, PR 86 863, PR 86 212, CJP 30 35, Nat 170 853, ArkF 5 427, Phil 43 648, Phil 43 221, Phca 17 637, PR 84 565, PR 83 680, PR 83 679, PR 82 277, PR 82 103, PR 81 642, PR 81 482, Nucl 7n 24, PR 76 94, PR 75 165, PR 75 1544, PR 74 1640)

\( \gamma(\theta) \)

ZP 244 332, NP 82 289, ArkF 23 49, PR 90 849

\( \beta(\theta) \)

PR 145 907, PR 79 728

\( \beta \) transverse polariz(\( \theta \)): YadF 5 1037, NuCo 5 942

\( \beta \) circular polariz(\( \theta \)): ZP 179 62

nucl align: PR 120 1777

0.150 level of \( ^{131}\text{I} \):

\[ t_{1/2} = \begin{array}{ll}
0.94 \text{d} & \text{ns, delay coinc (DUzb n4 24)} \\
0.95 \text{d} & \text{ns, delay coinc (PR 140 B536)} \\
0.76 \text{d} & \text{ns, delay coinc (NP A161 479)} \\
\end{array} \]

others (IzF 23 1445, ZETF 37 314, ArkF 11 10A, NP 1 821)

\[ \mu = 2.77 \text{d} \text{if } t_{1/2} = 0.95 \text{d} \text{ns, IPAC (NP A102 203)} \]

1.797 level of \( ^{131}\text{I} \):

\[ t_{1/2} = \begin{array}{ll}
5.9 \text{d} & \text{ns delay coinc (PR 140 B536)} \\
\end{array} \]

\[ \delta = -0.16 \text{d} \text{if } t_{1/2} = 5.9 \text{d} \text{ns, IPAC (NP A102 203)} \]

Fig. 5

XBL 7311-1442
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Atomic Energy Commission, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.