Brookhaven National Laboratory

U.S. DEPARTMENT of ENERGY

Unifying the ENDF free text documentation formats

David Brown
National Nuclear Data Center
dbrown@bnl.gov

7 Jan. 2025

@BrookhavenLab

Goal: Rewrite all ENDF documentation in common style

- Human readable
- Capture ‘Look and feel’ of well written documentation
- Structured, so machine readable
  - deterministic formatting
  - unified metadata
  - bibliography
  - within each reaction or class of reactions:
    - description
    - change log

Also, this is a chance to experiment with LLMs!

Brookhaven National Laboratory

Challenges

The expected ones

- Random formatting from past evaluators
- Full (or not) version history
- Programmatically generated text (EMPIRE & TALYS)
- File size
- ENDF vs. GNDS formatting
- Not really understanding nuclear physics

The unexpected ones

- Gosh 👏 darn 💩 emojis 🤬 and other non-ASCII characters
- The random seed & lack of predictability
- Sycophantic tone
- Occasionally ignoring instructions or missing content

I tried to follow best practices for prompt engineering (but I am listening to a sycophantic and stupid AI)

🧩 Summary

Practice Type	Example You Used	Quality	Suggestion
Context Setup	Role + domain	✅ Excellent	Add schema preamble
Multi-source Structuring	SOURCE 1–6	✅ Excellent	Maybe add internal cross-source references
Output Constraints	ASCII + table width	✅ Excellent	Add character validation script if automated
Conditional Logic	Include/omit rules	✅ Excellent	Consider flowchart annotation for automation
Validation Rules	Pre-output checklist	✅ Excellent	Add self-check assertions
Style Consistency	Golden example + YAML	✅ Excellent	Add section-specific micro-examples

I wrote a code to generate the prompt based on the ENDF/GNDS file content

Brc
Natio

I tried to follow best practices for prompt engineering (but I am listening to a sycophantic and stupid AI)

🧩 Summary

Practice Type	Example You Used	Quality	Suggestion
Context Setup	Role + domain	✅ Excellent	Add schema preamble
Multi-source Structuring	SOURCE 1–6	✅ Excellent	Maybe add internal cross-source references
Output Constraints	ASCII + table width	✅ Excellent	Add character validation script if automated
Conditional Logic	Include/omit rules	✅ Excellent	Consider flowchart annotation for automation
Validation Rules	Pre-output checklist	✅ Excellent	Add self-check assertions
Style Consistency	Golden example + YAML	✅ Excellent	Add section-specific micro-examples

My prompt generating code has gotten quite elaborate and I am finding my instruction inconsistency rate grows with the prompt size

I tried a variety of LLM's of different size & complexity, all using my superprompt

gpt-oss:20b and 120b

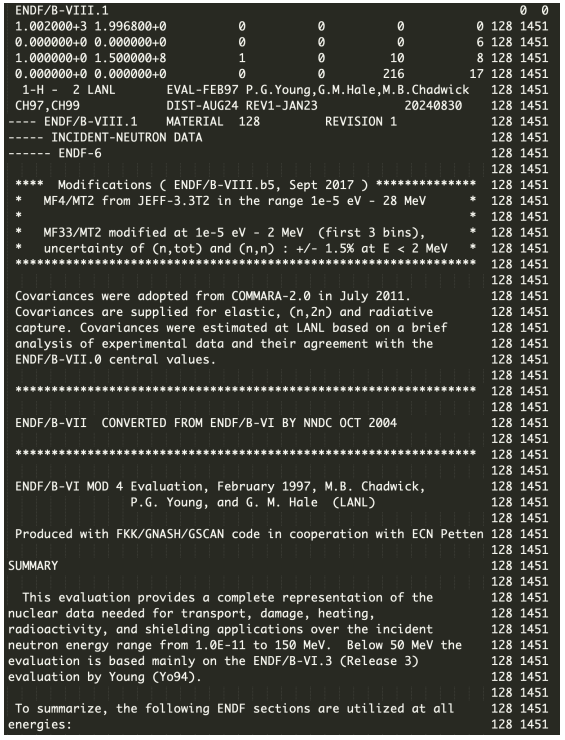
Qwen3's 32b-fp16, 32b-q8_0, 32b

Deepseek R1's 8b, 32b, 1.7b and 70b models

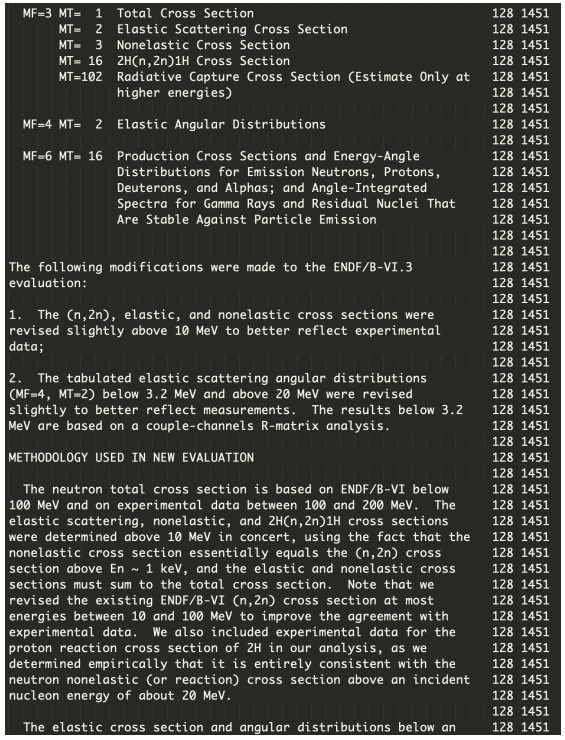
OpenAI's 3o, 4o and 5 (plus a few other varients)

=> 4o and 5 behaved best, but cost the most $$

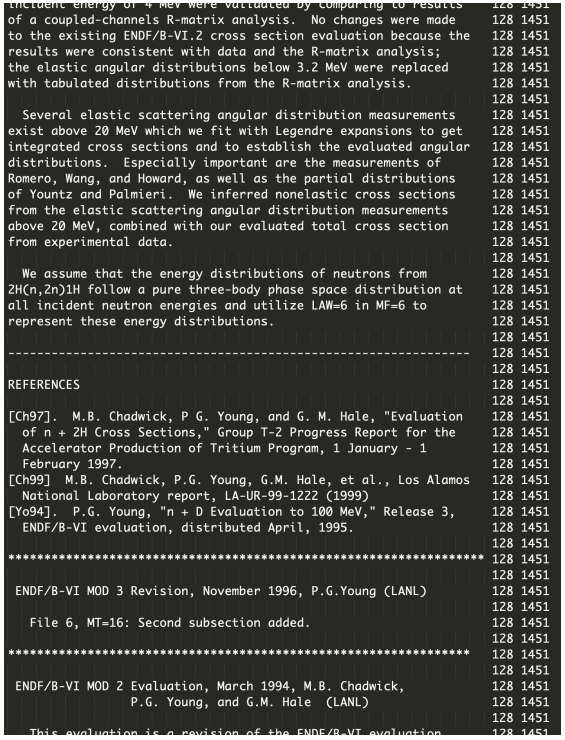
Deuterium (before)



```
 ENDF/B-VIII.1                                                          0  0
 1.002000+3 1.996800+0          0          0          0          0 128 1451
 0.000000+0 0.000000+0          0          0          0          6 128 1451
 1.000000+0 1.500000+8          1          0         10          8 128 1451
 0.000000+0 0.000000+0          0          0        216         17 128 1451
  1-H -  2 LANL       EVAL-FEB97 P.G.Young,G.M.Hale,M.B.Chadwick   128 1451
 CH97,CH99            DIST-AUG24 REV1-JAN23           20240830    128 1451
---- ENDF/B-VIII.1    MATERIAL  128         REVISION 1             128 1451
----- INCIDENT-NEUTRON DATA                                        128 1451
------ ENDF-6                                                      128 1451
                                                                   128 1451
 ****  Modifications ( ENDF/B-VIII.b5, Sept 2017 ) **************   128 1451
 *   MF4/MT2 from JEFF-3.3T2 in the range 1e-5 eV - 28 MeV      *   128 1451
 *                                                              *   128 1451
 *   MF33/MT2 modified at 1e-5 eV - 2 MeV  (first 3 bins),      *   128 1451
 *   uncertainty of (n,tot) and (n,n) : +/- 1.5% at E < 2 MeV   *   128 1451
 ************************************************************   128 1451
                                                                   128 1451
 Covariances were adopted from COMMARA-2.0 in July 2011.           128 1451
 Covariances are supplied for elastic, (n,2n) and radiative        128 1451
 capture. Covariances were estimated at LANL based on a brief      128 1451
 analysis of experimental data and their agreement with the        128 1451
 ENDF/B-VII.0 central values.                                      128 1451
                                                                   128 1451
 ************************************************************   128 1451
                                                                   128 1451
 ENDF/B-VII  CONVERTED FROM ENDF/B-VI BY NNDC OCT 2004             128 1451
                                                                   128 1451
 ************************************************************   128 1451
                                                                   128 1451
 ENDF/B-VI MOD 4 Evaluation, February 1997, M.B. Chadwick,         128 1451
                P.G. Young, and G. M. Hale  (LANL)                 128 1451
                                                                   128 1451
 Produced with FKK/GNASH/GSCAN code in cooperation with ECN Petten 128 1451
                                                                   128 1451
SUMMARY                                                            128 1451
                                                                   128 1451
  This evaluation provides a complete representation of the        128 1451
nuclear data needed for transport, damage, heating,                128 1451
radioactivity, and shielding applications over the incident        128 1451
neutron energy range from 1.0E-11 to 150 MeV.  Below 50 MeV the    128 1451
evaluation is based mainly on the ENDF/B-VI.3 (Release 3)          128 1451
evaluation by Young (Yo94).                                        128 1451
                                                                   128 1451
 To summarize, the following ENDF sections are utilized at all     128 1451
energies:                                                          128 1451
```



```
  MF=3 MT=  1  Total Cross Section                                 128 1451
       MT=  2  Elastic Scattering Cross Section                    128 1451
       MT=  3  Nonelastic Cross Section                            128 1451
       MT= 16  2H(n,2n)1H Cross Section                            128 1451
       MT=102  Radiative Capture Cross Section (Estimate Only at   128 1451
               higher energies)                                    128 1451
                                                                   128 1451
  MF=4 MT=  2  Elastic Angular Distributions                       128 1451
                                                                   128 1451
  MF=6 MT= 16  Production Cross Sections and Energy-Angle          128 1451
               Distributions for Emission Neutrons, Protons,       128 1451
               Deuterons, and Alphas; and Angle-Integrated         128 1451
               Spectra for Gamma Rays and Residual Nuclei That     128 1451
               Are Stable Against Particle Emission                128 1451
                                                                   128 1451
                                                                   128 1451
The following modifications were made to the ENDF/B-VI.3           128 1451
evaluation:                                                        128 1451
                                                                   128 1451
1.  The (n,2n), elastic, and nonelastic cross sections were        128 1451
revised slightly above 10 MeV to better reflect experimental       128 1451
data;                                                              128 1451
                                                                   128 1451
2.  The tabulated elastic scattering angular distributions         128 1451
(MF=4, MT=2) below 3.2 MeV and above 20 MeV were revised           128 1451
slightly to better reflect measurements.  The results below 3.2    128 1451
MeV are based on a couple-channels R-matrix analysis.              128 1451
                                                                   128 1451
METHODOLOGY USED IN NEW EVALUATION                                 128 1451
                                                                   128 1451
  The neutron total cross section is based on ENDF/B-VI below      128 1451
100 MeV and on experimental data between 100 and 200 MeV.  The     128 1451
elastic scattering, nonelastic, and 2H(n,2n)1H cross sections      128 1451
were determined above 10 MeV in concert, using the fact that the   128 1451
nonelastic cross section essentially equals the (n,2n) cross       128 1451
section above En ~ 1 keV, and the elastic and nonelastic cross     128 1451
sections must sum to the total cross section.  Note that we        128 1451
revised the existing ENDF/B-VI (n,2n) cross section at most        128 1451
energies between 10 and 100 MeV to improve the agreement with      128 1451
experimental data.  We also included experimental data for the     128 1451
proton reaction cross section of 2H in our analysis, as we         128 1451
determined empirically that it is entirely consistent with the     128 1451
neutron nonelastic (or reaction) cross section above an incident   128 1451
nucleon energy of about 20 MeV.                                    128 1451
                                                                   128 1451
  The elastic cross section and angular distributions below an     128 1451
```



```
incident energy of 4 MeV were validated by comparing to results    128 1451
of a coupled-channels R-matrix analysis.  No changes were made     128 1451
to the existing ENDF/B-VI.2 cross section evaluation because the   128 1451
results were consistent with data and the R-matrix analysis;       128 1451
the elastic angular distributions below 3.2 MeV were replaced      128 1451
with tabulated distributions from the R-matrix analysis.           128 1451
                                                                   128 1451
  Several elastic scattering angular distribution measurements    128 1451
exist above 20 MeV which we fit with Legendre expansions to get    128 1451
integrated cross sections and to establish the evaluated angular   128 1451
distributions.  Especially important are the measurements of       128 1451
Romero, Wang, and Howard, as well as the partial distributions     128 1451
of Yountz and Palmieri.  We inferred nonelastic cross sections     128 1451
from the elastic scattering angular distribution measurements      128 1451
above 20 MeV, combined with our evaluated total cross section      128 1451
from experimental data.                                            128 1451
                                                                   128 1451
  We assume that the energy distributions of neutrons from         128 1451
2H(n,2n)1H follow a pure three-body phase space distribution at    128 1451
all incident neutron energies and utilize LAW=6 in MF=6 to         128 1451
represent these energy distributions.                              128 1451
                                                                   128 1451
-------------------------------------------------------------      128 1451
                                                                   128 1451
REFERENCES                                                         128 1451
                                                                   128 1451
[Ch97].  M.B. Chadwick, P G. Young, and G. M. Hale, "Evaluation    128 1451
  of n + 2H Cross Sections," Group T-2 Progress Report for the     128 1451
  Accelerator Production of Tritium Program, 1 January - 1         128 1451
  February 1997.                                                   128 1451
[Ch99]  M.B. Chadwick, P.G. Young, G.M. Hale, et al., Los Alamos   128 1451
  National Laboratory report, LA-UR-99-1222 (1999)                 128 1451
[Yo94].  P.G. Young, "n + D Evaluation to 100 MeV," Release 3,     128 1451
  ENDF/B-VI evaluation, distributed April, 1995.                   128 1451
                                                                   128 1451
************************************************************       128 1451
                                                                   128 1451
 ENDF/B-VI MOD 3 Revision, November 1996, P.G.Young (LANL)         128 1451
                                                                   128 1451
   File 6, MT=16: Second subsection added.                         128 1451
                                                                   128 1451
************************************************************       128 1451
                                                                   128 1451
 ENDF/B-VI MOD 2 Evaluation, March 1994, M.B. Chadwick,            128 1451
                P.G. Young, and G.M. Hale  (LANL)                  128 1451
                                                                   128 1451
 This evaluation is a revision of the ENDF/B-VI evaluation         128 1451
```

+ 2 more revisions, with very different formatting

Brookhaven National Laboratory

Deuterium (after)

```
                                                                   1 0  0
 1.002000+3 1.996800+0          0          0          0          0 128 1451
 0.000000+0 0.000000+0          0          0          0          6 128 1451
 1.000000+0 1.500000+8          1          0         10          8 128 1451
 0.000000+0 0.000000+0          0          0        164         16 128 1451
  1-H -  2 LANL       EVAL-FEB97 P.G.Young,G.M.Hale,M.B.Chadwick   128 1451
 CH97,CH99            DIST-AUG24 REV1-JAN23           20240830    128 1451
---- ENDF/B-VIII.1    MATERIAL  128         REVISION 1             128 1451
----- INCIDENT-NEUTRON DATA                                       128 1451
------ ENDF-6                                                     128 1451
                                                                  128 1451
n + 2H Evaluation (ENDF/B-VIII.1 Rev.1)                           128 1451
=====================================                             128 1451
:Authors: P.G. Young, G.M. Hale, M.B. Chadwick                    128 1451
:Laboratory: LANL                                                 128 1451
:Evaluation date: FEB97                                           128 1451
:Projectile: n                                                    128 1451
:Target: H2                                                       128 1451
:Compound nucleus: H3                                             128 1451
:Projectile frame: lab                                            128 1451
:Temperature: 0. K                                                128 1451
:ENDF release: ENDF/B-VIII.1 Rev.1                                128 1451
:GNDS version: 2.0                                                128 1451
                                                                  128 1451
..  This documentation was summarized and formatted by chatGPT    128 1451
version 4o and verified by David A. Brown (dbrown@bnl.gov).       128 1451
                                                                  128 1451
Introduction                                                      128 1451
------------                                                      128 1451
This evaluation provides a complete representation of the nuclear 128 1451
data needed for transport, damage, heating, radioactivity, and    128 1451
shielding applications over the incident neutron energy range from 128 1451
1.0E-11 to 150 MeV. Below 50 MeV, the evaluation is based mainly  128 1451
on the ENDF/B-VI.3 (Release 3) evaluation by Young [3]. This      128 1451
release supersedes ENDF/B-VI.3.                                   128 1451
                                                                  128 1451
Methodology Used in New Evaluation                                128 1451
----------------------------------                                128 1451
The neutron total cross section is based on ENDF/B-VI below 100   128 1451
MeV and on experimental data between 100-200 MeV. Above 10 MeV,   128 1451
the elastic, nonelastic, and (n,2n) cross sections were determined 128 1451
in concert, using the fact that the nonelastic cross section      128 1451
approximately equals the (n,2n) cross section above En ~ 1 keV.   128 1451
The total cross section is partitioned such that elastic plus     128 1451
nonelastic equals total.                                          128 1451
                                                                  128 1451
```

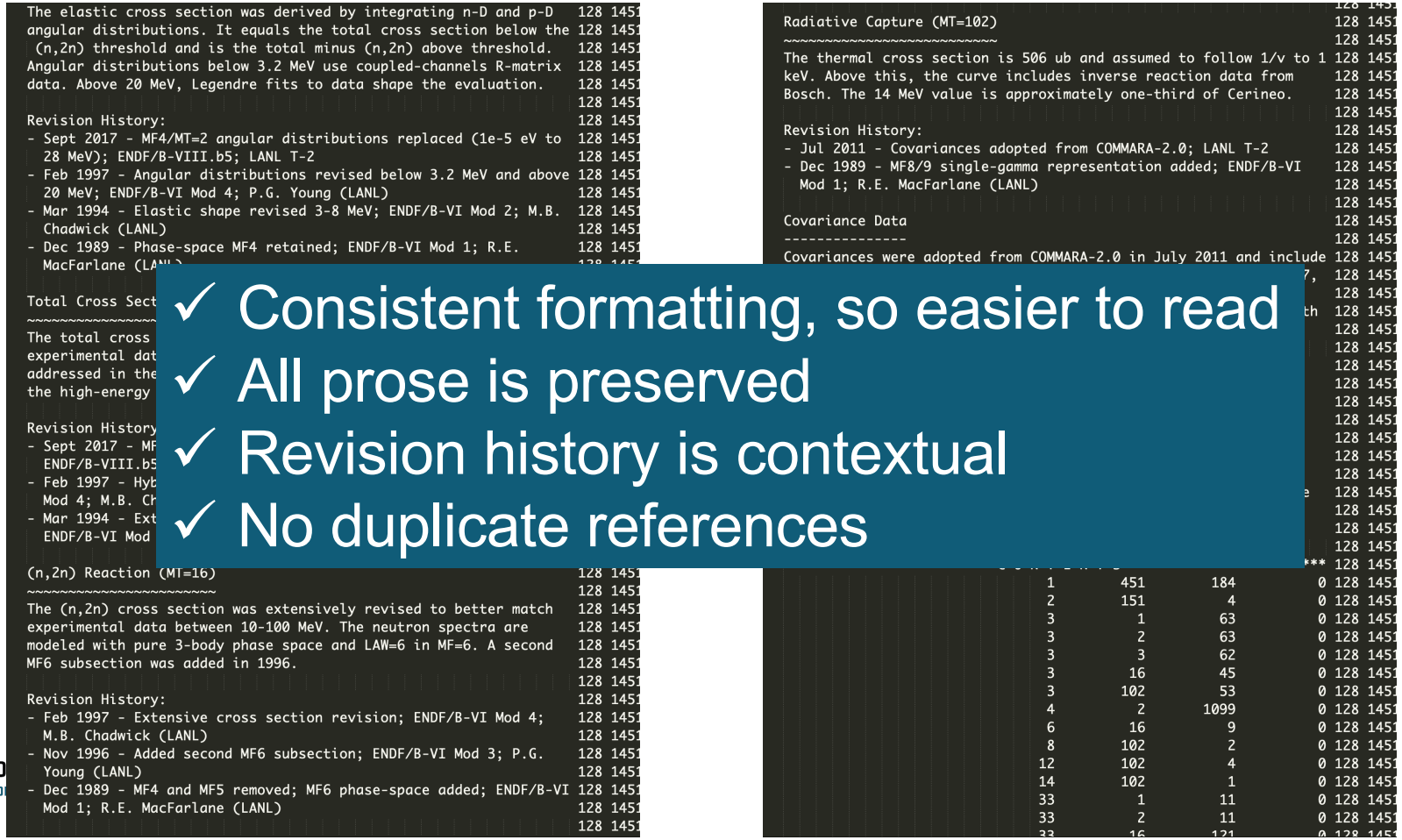
```
Revisions were made to the ENDF/B-VI (n,2n) cross section at most  128 1451
energies between 10-100 MeV to improve agreement with experimental 128 1451
data. Proton reaction cross section data on 2H were included and   128 1451
found consistent with neutron nonelastic cross sections above 20   128 1451
MeV.                                                               128 1451
                                                                   128 1451
Below 4 MeV, the elastic cross section and angular distributions   128 1451
were validated using a coupled-channels R-matrix analysis. While   128 1451
the ENDF/B-VI.2 cross section was retained, angular distributions  128 1451
below 3.2 MeV were replaced with tabulated R-matrix results.       128 1451
                                                                   128 1451
Above 20 MeV, angular distributions were based on Legendre fits to 128 1451
measurements by Romero, Wang, Howard, Yountz, and Palmieri.        128 1451
Nonelastic cross sections were inferred from elastic angular       128 1451
measurements and the evaluated total cross section.                128 1451
                                                                   128 1451
The emitted neutron spectra from (n,2n) reactions were modeled as  128 1451
pure three-body phase space (LAW=6 in MF=6) at all incident        128 1451
energies.                                                          128 1451
                                                                   128 1451
Neutron Resonance Region                                           128 1451
------------------------                                           128 1451
No resolved resonance parameters (MF=2) are provided in this       128 1451
evaluation. However, R-matrix analysis was used to validate low-   128 1451
energy elastic behavior. The following table summarizes integral   128 1451
metrics relevant to the 1/v region:                                128 1451
+----------+---------+---------+---------+----------+              128 1451
| Reaction |    RI   | 2200 m/s| Westcott|   MACS   |              128 1451
|          |   [b]   |   [b]   | factor  | 30keV[mb]|              128 1451
+==========+=========+=========+=========+==========+              128 1451
| Elastic  | 5.366e+1| 3.395e+0| 1.1284  | 3.660e+3 |              128 1451
| Capture  | 2.753e-4| 5.064e-4| 0.8159  | 1.999e-3 |              128 1451
| Total    | 5.406e+1| 3.396e+0| 1.1285  | 3.660e+3 |              128 1451
+----------+---------+---------+---------+----------+              128 1451
                                                                   128 1451
Revision History:                                                  128 1451
- Sept 2017 - MF33 uncertainties propagated to resonance metrics;  128 1451
  ENDF/B-VIII.b5; LANL T-2                                         128 1451
- Feb 1997 - R-matrix analysis validated elastic RI; ENDF/B-VI     128 1451
  Mod 4; P.G. Young (LANL)                                         128 1451
                                                                   128 1451
Reaction Summaries                                                 128 1451
------------------                                                 128 1451
                                                                   128 1451
Elastic Scattering (MT=2)                                          128 1451
~~~~~~~~~~~~~~~~~~~~~~~                                            128 1451
```

Deuterium (after)

```
The elastic cross section was derived by integrating n-D and p-D   128 1451
angular distributions. It equals the total cross section below the 128 1451
 (n,2n) threshold and is the total minus (n,2n) above threshold.   128 1451
Angular distributions below 3.2 MeV use coupled-channels R-matrix  128 1451
data. Above 20 MeV, Legendre fits to data shape the evaluation.    128 1451
                                                                   128 1451
Revision History:                                                  128 1451
- Sept 2017 - MF4/MT=2 angular distributions replaced (1e-5 eV to  128 1451
  28 MeV); ENDF/B-VIII.b5; LANL T-2                                128 1451
- Feb 1997 - Angular distributions revised below 3.2 MeV and above 128 1451
  20 MeV; ENDF/B-VI Mod 4; P.G. Young (LANL)                       128 1451
- Mar 1994 - Elastic shape revised 3-8 MeV; ENDF/B-VI Mod 2; M.B.  128 1451
  Chadwick (LANL)                                                  128 1451
- Dec 1989 - Phase-space MF4 retained; ENDF/B-VI Mod 1; R.E.       128 1451
  MacFarlane (LANL)                                                128 1451
                                                                   128 1451
Total Cross Section (MT=1)                                         128 1451
~~~~~~~~~~~~~~~~~~~~~~~~~~                                         128 1451
The total cross section uses ENDF/B-VI data below 100 MeV and      128 1451
experimental data above. Discrepancies from LA-3271 [4] were       128 1451
addressed in the 3-8 MeV range. Proton-induced data corroborated   128 1451
the high-energy evaluation.                                        128 1451
                                                                   128 1451
Revision History:                                                  128 1451
- Sept 2017 - MF33 uncertainties for MT=1 tightened below 2 MeV;   128 1451
  ENDF/B-VIII.b5; LANL T-2                                         128 1451
- Feb 1997 - Hybrid experimental approach above 10 MeV; ENDF/B-VI  128 1451
  Mod 4; M.B. Chadwick (LANL)                                      128 1451
- Mar 1994 - Extended to 100 MeV and shape updated 3-8 MeV;        128 1451
  ENDF/B-VI Mod 2; M.B. Chadwick (LANL)                            128 1451
                                                                   128 1451
(n,2n) Reaction (MT=16)                                            128 1451
~~~~~~~~~~~~~~~~~~~~~~~                                            128 1451
The (n,2n) cross section was extensively revised to better match   128 1451
experimental data between 10-100 MeV. The neutron spectra are      128 1451
modeled with pure 3-body phase space and LAW=6 in MF=6. A second   128 1451
MF6 subsection was added in 1996.                                  128 1451
                                                                   128 1451
Revision History:                                                  128 1451
- Feb 1997 - Extensive cross section revision; ENDF/B-VI Mod 4;    128 1451
  M.B. Chadwick (LANL)                                             128 1451
- Nov 1996 - Added second MF6 subsection; ENDF/B-VI Mod 3; P.G.    128 1451
  Young (LANL)                                                     128 1451
- Dec 1989 - MF4 and MF5 removed; MF6 phase-space added; ENDF/B-VI 128 1451
  Mod 1; R.E. MacFarlane (LANL)                                    128 1451
                                                                   128 1451
```

```
                                                                   128 1451
Radiative Capture (MT=102)                                         128 1451
~~~~~~~~~~~~~~~~~~~~~~~~~~                                         128 1451
The thermal cross section is 506 ub and assumed to follow 1/v to 1 128 1451
keV. Above this, the curve includes inverse reaction data from     128 1451
Bosch. The 14 MeV value is approximately one-third of Cerineo.     128 1451
                                                                   128 1451
Revision History:                                                  128 1451
- Jul 2011 - Covariances adopted from COMMARA-2.0; LANL T-2        128 1451
- Dec 1989 - MF8/9 single-gamma representation added; ENDF/B-VI    128 1451
  Mod 1; R.E. MacFarlane (LANL)                                    128 1451
                                                                   128 1451
Covariance Data                                                    128 1451
---------------                                                    128 1451
Covariances were adopted from COMMARA-2.0 in July 2011 and include 128 1451
MT=2 (elastic), MT=16 (n,2n), and MT=102 (capture). In Sept 2017,  128 1451
uncertainties for MT=1 and MT=2 below 2 MeV were tightened to      128 1451
+/-1.5%, reflecting updates from LANL T-2 based on agreement with  128 1451
ENDF/B-VII.0 central values.                                       128 1451
                                                                   128 1451
References                                                         128 1451
----------                                                         128 1451
[1] M.B. Chadwick, P.G. Young, G.M. Hale, "Evaluation of n + 2H    128 1451
    Cross Sections," Group T-2 Progress Report for the APT         128 1451
    Program, 1 Jan - 1 Feb 1997.                                   128 1451
[2] M.B. Chadwick, P.G. Young, G.M. Hale, et al., LA-UR-99-1222    128 1451
    (1999).                                                        128 1451
[3] P.G. Young, "n + D Evaluation to 100 MeV," ENDF/B-VI Release   128 1451
    3, Apr 1995.                                                   128 1451
[4] D.H. Stewart & R.F. Horsley, LA-3271, 1968.                    128 1451
                                                                   128 1451
 ********************** C O N T E N T S ********************** 128 1451
                                1       451       184          0 128 1451
                                2       151         4          0 128 1451
                                3         1        63          0 128 1451
                                3         2        63          0 128 1451
                                3         3        62          0 128 1451
                                3        16        45          0 128 1451
                                3       102        53          0 128 1451
                                4         2      1099          0 128 1451
                                6        16         9          0 128 1451
                                8       102         2          0 128 1451
                               12       102         4          0 128 1451
                               14       102         1          0 128 1451
                               33         1        11          0 128 1451
                               33         2        11          0 128 1451
                               33        16       121          0 128 1451
```

Deuterium (after)



- ✓ Consistent formatting, so easier to read
- ✓ All prose is preserved
- ✓ Revision history is contextual
- ✓ No duplicate references

How did it go so far?

- n+1H – worked OK, but inconsistent performance
- n+2H - very simple documentation, formatting worked very well
- n+3H - very simple documentation, formatting worked very well
- n+3He – bigger file, got confused occasionally
- n+4He – very simple documentation, formatting worked very well
- n+16O – too big, went badly

ChatGPT 5 recommendation

7. Chunked Execution Instruction

If you plan to feed this to an LLM via an API, instruct:

> "Process and emit sections incrementally (Metadata, Resonance Region, Reactions, etc.), verifying each before proceeding."

This prevents truncation and maintains compliance in long outputs — essentially **streamed prompt execution**.

If you want to try it out or improve it, please do!

https://git.nndc.bnl.gov/endf/tools/doc-fixer.git