Modernization of the Evaluated Nuclear Structure Data File (ENSDF)

A. Hayes, E.A. McCutchan, B. Shu, A. Sonzogni, S. Zhu, A. Mattera – NNDC, BNL
S. Yoo, C. Soto, S. McCorkle – CSI, BNL
F. Kondev - ANL
C. Mattoon – LLNL

BROOKHAVEN NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

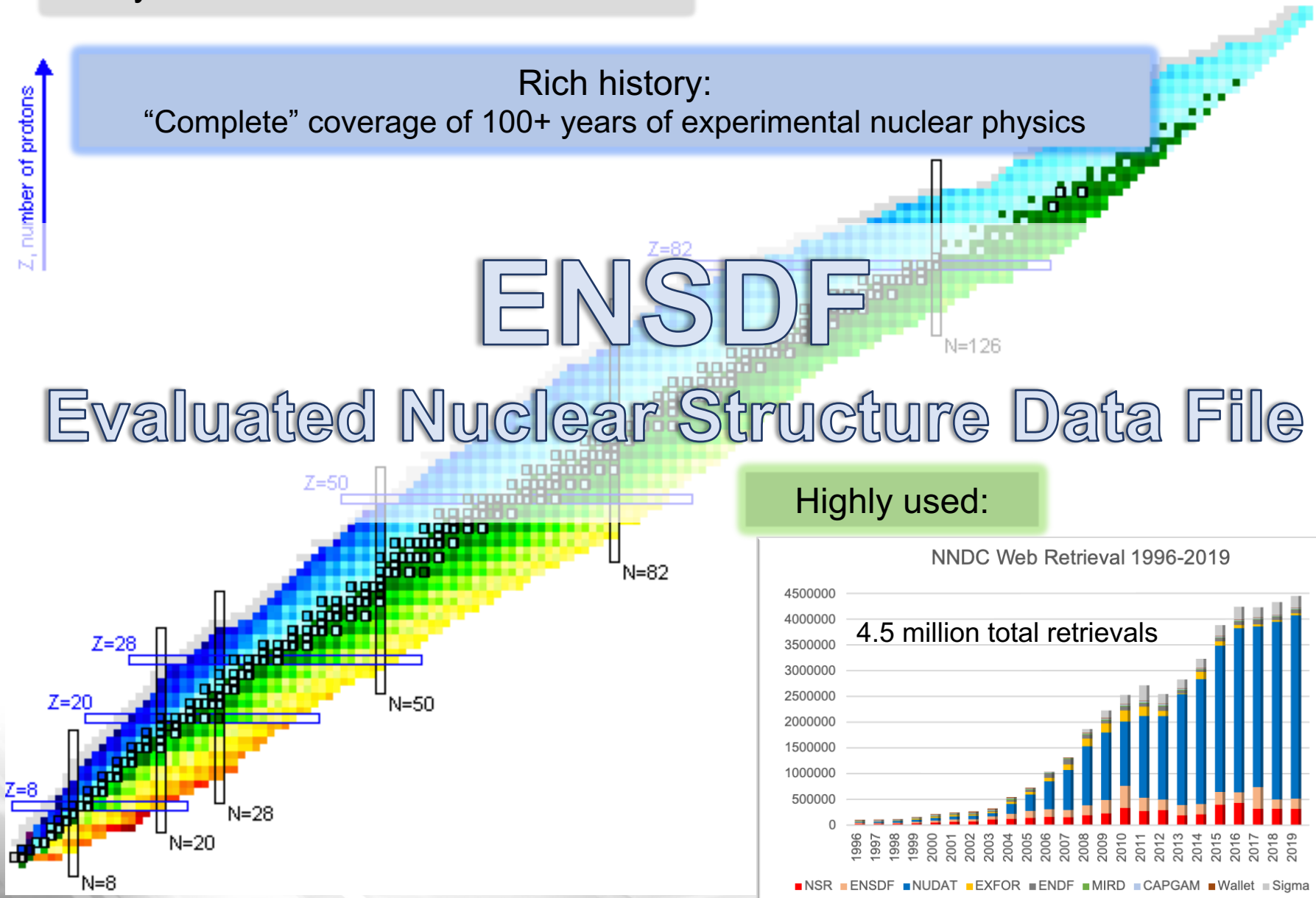
BROOKHAVEN SCIENCE ASSOCIATES

Unique :
Only database of its kind in the world

Rich history:
"Complete" coverage of 100+ years of experimental nuclear physics

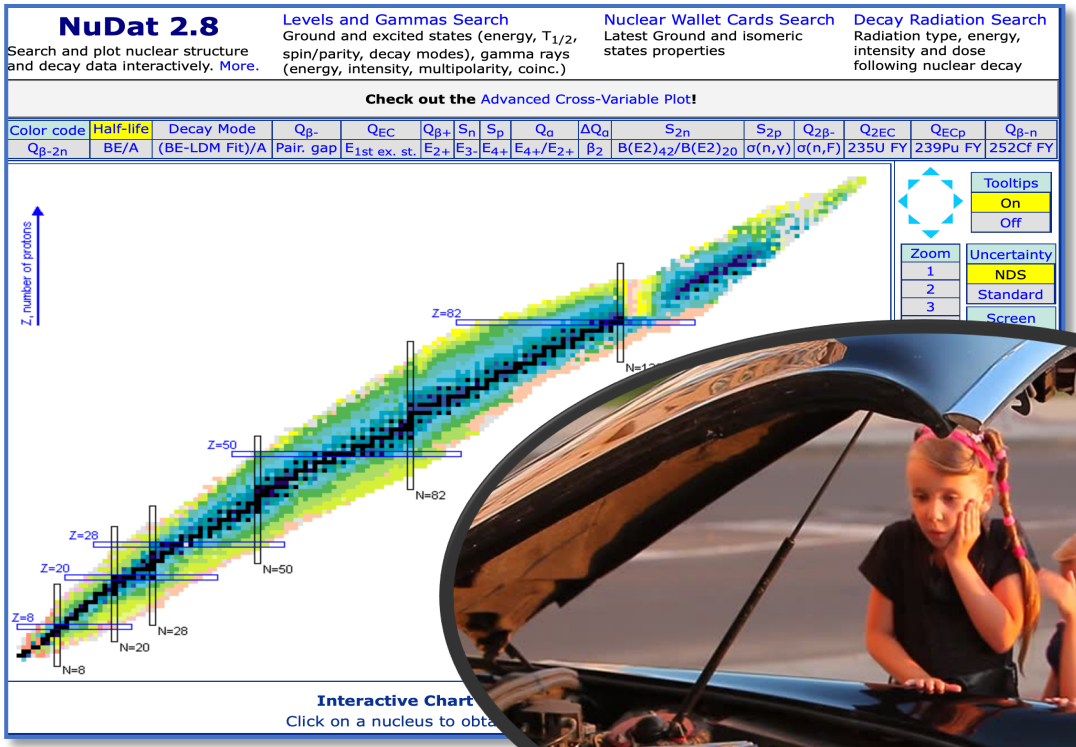
ENSDF

Evaluated Nuclear Structure Data File



Highly used:

The faces of ENSDF



NuDat – Graphical and tabular interface

ENSDF web app – Tables and drawings in the “Nuclear Data Sheets” style

				K	^{149}Sm(p,t)
				L	^{151}Eu(μ^-,4nγ)
				M	Coulomb excitation
121.212[&] *5*	5/2$^-$	0.798 ns *17*	AB DEFG I	M	μ=−0.449 *25* (1989Ra17); Q=−0.45 *19* (1989Ra17)

Comments

%α=100

μ=−0.8148 *7* (2005St24); Q=−0.261 *7* (2005St24)

μ: measured by atomic beam magnetic resonance – thermal beam (1966Wo05).

Q: measured by atomic beam magnetic resonance – thermal beam (re-evaluated data) (1992Le09).

J$^\pi$: from 1976Fu06, π from L(p,t)=0.

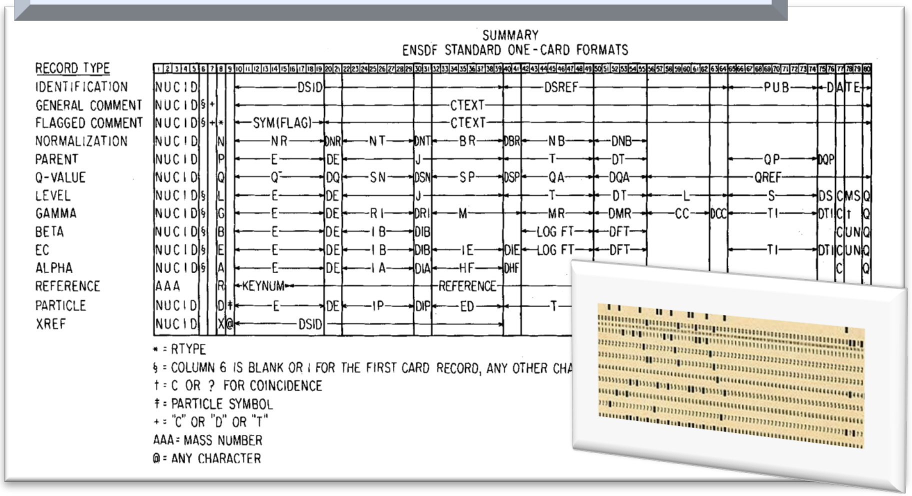
rms charge radius: 4.9839 *10* (2004An14).

μ,Q: measured by Mossbauer effect (1971Pa04); the values given by 2005St24 are the rounded-off values of 1989Ra17 (μ=−0.45 *3*, Q=−0.5 *2*).

T$_{1/2}$: weighted av. of 0.80 ns *4* (1968Bo47), 0.78 ns *3* (1970Ko38), 0.77 ns *4* (1971Be53), 0.83 ns *3* (1978VyZV), in ^{147}Eu ε decay; 0.79 ns *14* from B(E2)↑ in Coulomb excitation.

Under the hood of ENSDF

80 Column ASCII format



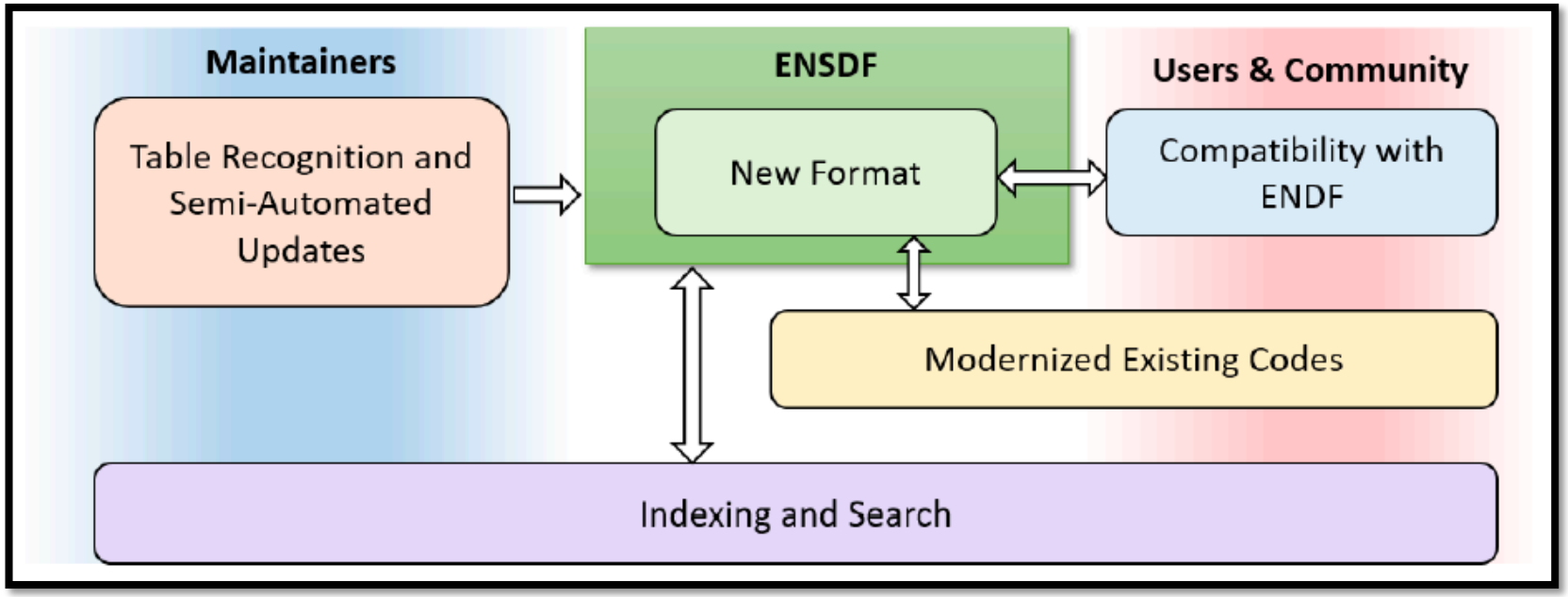
- All of nuclear physics doesn't fit in 80 columns
- No room to grow
- ML next to impossible
- Hard to engage next gen

```
137CS PN                                                                     6
137CS  L 0.0         7/2+                 30.08 Y    9                      A
137CSX L XREF=ACDEFGH
137CS2 L %B-=100$MOMM1=+2.8413 1 (1989Ra17)$MOME2=+0.051 1 (1989Ra17)
137CS cL T$Deduced by evaluators using the Limitation of Relative Statistical
137CS2cL Weights (LRSW) method for analyzing the following set of
137CS3cL discrepant (|h{+2}/|n=18.6) experimental values: 10970 d {I20}
137CS4cL (2004Sc04); 11018 d {I10} (2002Un02); 10941 d {I7} (1992Go24);
137CS5cL 10968 d {I5} (1990Ma15); 11009 d {I11} (1980Ho17); 10906 d {I33}
137CS6cL (1978Gr08); 11034 d {I29} (1973Co39); 11021 d {I5} (1973Di01); 11023 d
137CS7cL {I37} (1972Em01); 10921 d {I17} (1970Wa19); 11191 d {I157} (1970Ha32);
137CS8cL 11286 d {I256}, 10921 d {I183} (1965Fl01); 11220 d {I47} (1965Le25);
137CS9cL 10665 d {I110} (1963Ri02); 1
137CSAcL (1962Fl09); 11103 d {I146} (
137CSxcL 9715 d {I146} (1955Wi21). [1
137CS cL          Other evaluated re
137CS2cL 30.08 y {I3} (1996ChZY, 1994
```

A LOT of data stored in comments
Non-standardized entry

ENSDF Modernization

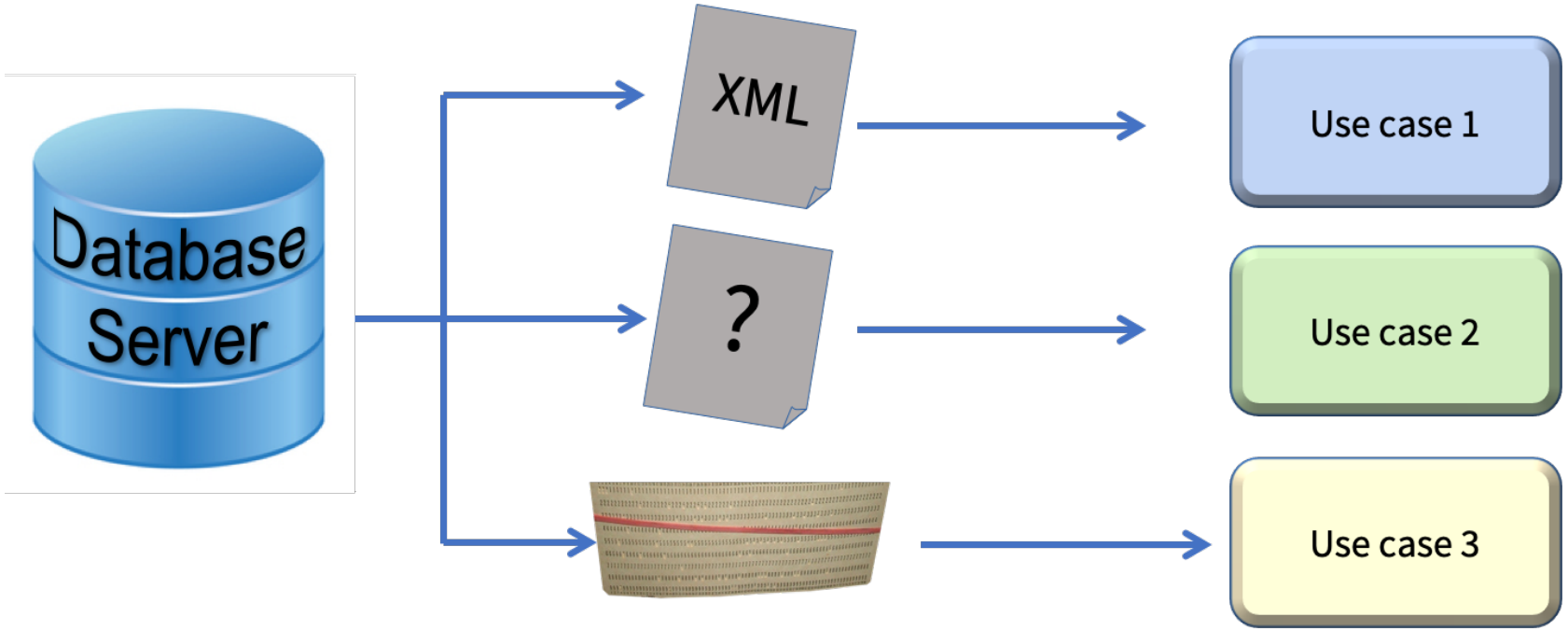
Funded through Nuclear Data Interagency Working Group – FOA LAB 19-2114
Collaboration with BNL (NNDC and CSI), ANL, LLNL – 3 year proposal



1) Develop a new Object-Oriented Database for ENSDF
2) Modernize existing codes used with ENSDF
3) Develop streamlined publication to ENSDF software
4) Implement indexing and searching
5) Ensure compatibility with end users

An object-oriented ENSDF database

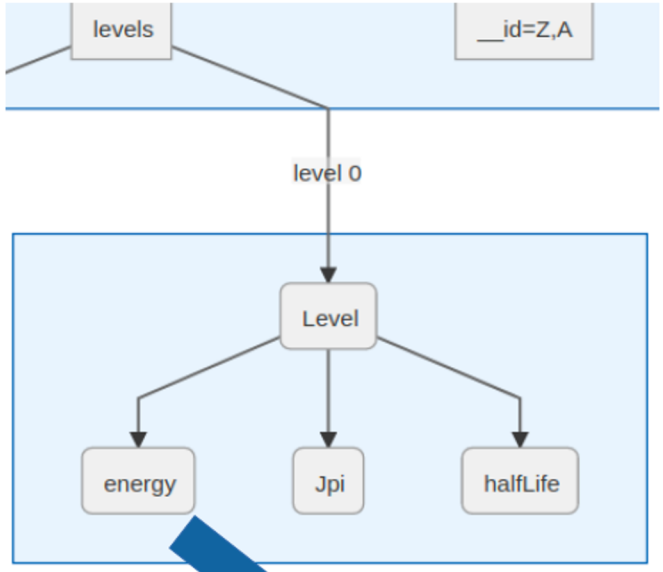
Think in terms of **databases** for **storage**;
Think in terms of **file formats** for **transmission.**



- Users of data can be ***agnostic*** about the database structure
- Database can be changed ***independently*** of the file format
- Database can store ***many data types;*** images, pdfs, experimental data …

We are moving forward with CouchDB

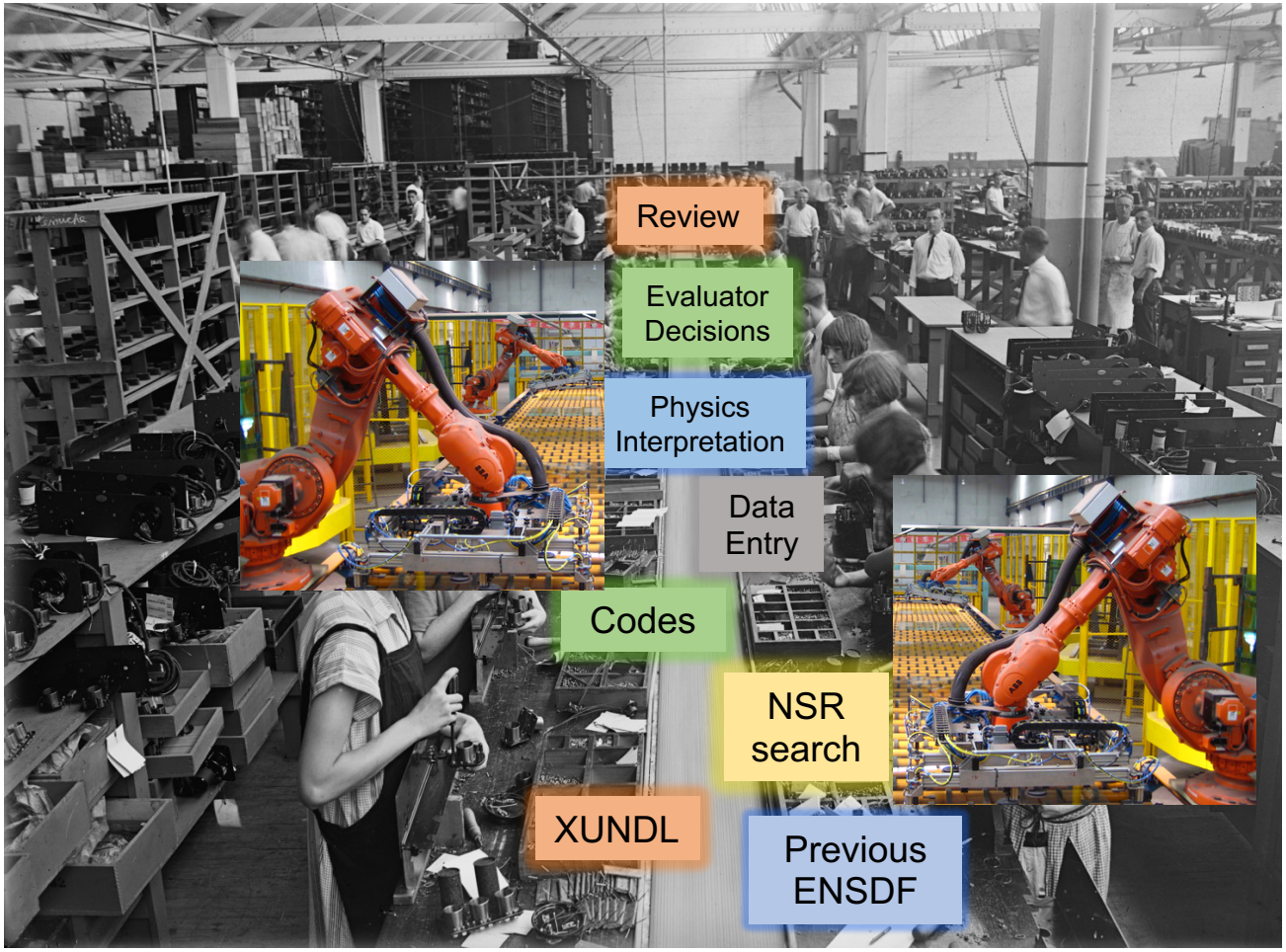
Year 1 : Format Development




```
"levels": {
    "0": {↔},
    "1": {
        "levelEnergy": {
            "energy": 100.0,
            "dEnergy": [-0.5, 1.0]
        },
        "halfLife": {},
```

list

Improving efficiency



- Many steps in evaluation can be automated
- Increases efficiency and reliability
- Evaluators can focus on physics and data science

Machine Learning to improve data compilation

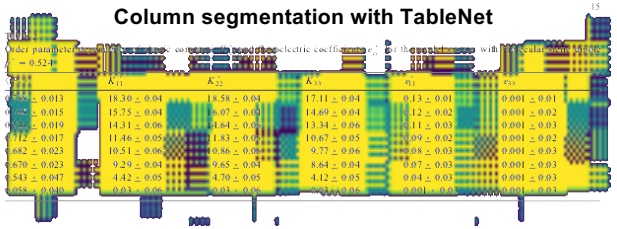
The Problem:

- Majority of data resides in tables
- Standard table extraction techniques are either highly manual, extremely noisy, or both

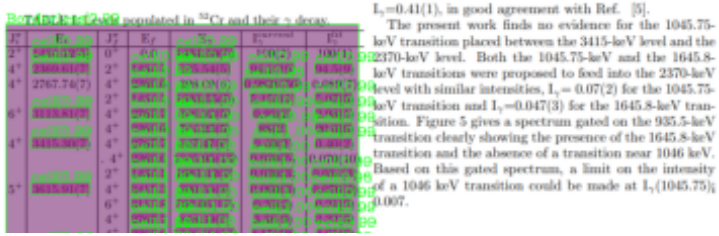
Our Solution:

- Isolate tables then extract contents with deep neural network
- Apply and then improve on visual AI-based segmentation methods

Column segmentation with TableNet



Borderless table segmentation and cell detection with CascadeTabNet



Noisy extraction results with Tabula (after manual alignment)

Jπ	E level	Jf π	Eγ	Iγ	σ	B(E2)lit a		B(E2)present B(E2)IBA
	(keV)		(keV)			(W.u.)		(W.u.) (W.u.)
2+1	355.6	0+1	355.6	106	+	40.6(2)		40.6(2)b 40.6
		2+1	1047.1	73(6)	- > 0.75 and < 5		d	2.5(11) 0
6+1	1525.8	4+1	649.4	96(8)	+	73+4−7		69(6) 53
2+ 4	1604.5	03 +	201.8	2.6(9)c	+			148(40) 22
		3+1	589.4	2.6(9)c	-			

Broken super/subscripts Wrong cell placement Merged columns

- Funded effort through NDIAWG for specific ND database
- Automatic table extraction would enable new, large-scale document analysis and knowledge discovery tools for many fields

We need input !!!

- Workshops for input and feedback
  - ENSDF Technical Meeting – April 2021
  - Workshop at LECM – August 2021
  - “Booth” at INNM - Summer 2021

- Surveys to ensure needs are encapsulated
  - General surveys posted to website - soon
  - Targeted surveys via email – soon

- Formation of ENSDF resource group
  - Volunteers to advise on a regular basis
  - Makeup from all ENSDF user communities

Community input now defines next 50 years of ENSDF !!