Purpose of the new data structure: Dave's Perspective

David Brown

Legacy formats

Will we continue to be enslaved by this "modern technology"?

Slide based on slide from M. Herman

Wednesday, November 28, 12
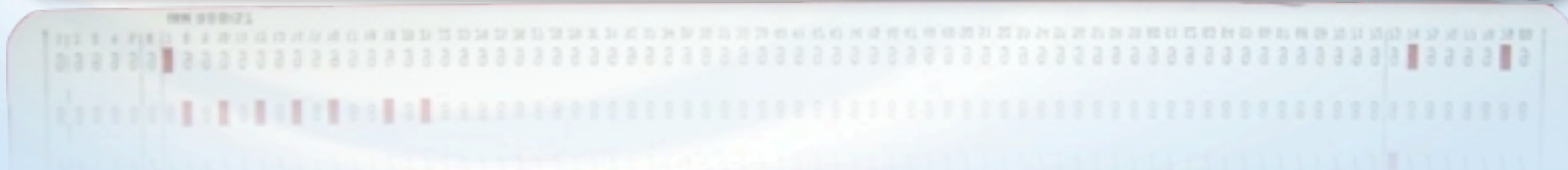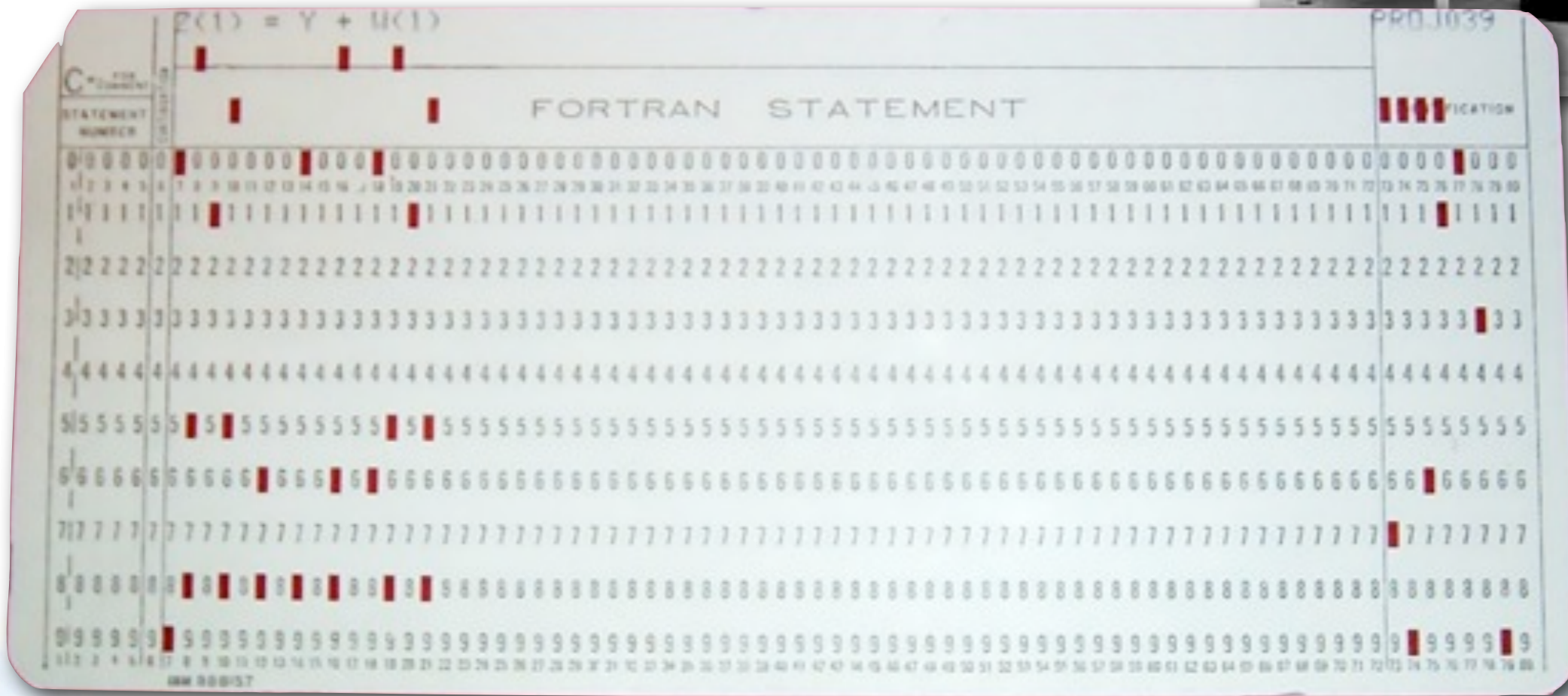
No, seriously

- A good format can determine the data structures used to interact with it
- These data structures are the components we use to create new things
- **We are trying to create a development environment (tools + components) that we enjoy working with**
- We will be working with these tools for a long time

Good tools == Happy developers

BROOKHAVEN
NATIONAL LABORATORY

Wednesday, November 28, 12

Our users determine the products we develop, so who are our users? Everybody...

Nuclear Science Community

✦ experiments

✦ theory

DANCE detector LANSCE

Nuclear Data Community

✦ compilation

✦ evaluation

✦ dissemination

✦ archival

Application Community

needs data:

✦ complete

✦ organized

✦ traceable

✦ readable

What makes an enjoyable development environment for nuclear data applications?

- **Simple**
  - Structures "obvious" to nuclear physicist/engineer
  - Don't want to have to read the manual to get something done
  - Enable rapid development
  - Enable good bookkeeping
- **Legacy support**
  - Don't like rewriting it all
  - Need something to benchmark against
- **Unrestricted**
  - Programming language agnostic
  - Unicode, localization support
  - Don't box us in with poor design choices

- Open source
- No export controls
- **Reusable**
  - Write once, reuse often
  - Tested, trusted, discrete components
- **Powerful**
  - Smartphone to Supercomputer
  - Big iron, Big data
  - All nuclei in the *Table of Isotopes* and all the particles in the *The Review of Particle Physics*
  - From thermal neutrons to GeV's and beyond

What makes an enjoyable development environment for nuclear data applications?

- **Simple**
  - Structures "obvious" to nuclear physicist/engineer
  - Don't want to have to read the manual to get something done
  - Enable rapid development
  - Enable good bookke...

- **Legacy support**
  - Don't like rewriting it all
  - Need something to benchmark against

- **Unrestricted**
  - Programming language agnostic
  - Unicode, localization support
  - Don't box us in with poor design choices

- Open source
- No export controls

- **Reusable**
  - Write once, reuse often
  - Tested, trusted, discrete components

- **...ul**
  - ...hone to Supercomputer
  - Big iron, Big data
  - All nuclei in the *Table of Isotopes* and all the particles in the *The Review of Particle Physics*
  - From thermal neutrons to GeV's and beyond

S.L.U.R.P.

BROOKHAVEN
NATIONAL LABORATORY

Wednesday, November 28, 12

We as a community support many database products

Library	Root node	Main use
NSR	Publication	**Bibliographic**: archival
EXFOR	Publication	**Experimental data**: archival; reaction evaluation; basic science
ENSDF, XUNDL	Nucleus	**Structure data**: archival, reaction evaluation, basic science
RIPL	Parameter type	**Input parameters**: reaction evaluation
ENDF	Target+Projectile	**Reaction data**: particle transport, activation

The hierarchy of a library should be clear from the application domain, *without* relying on documentation

EXFOR format already is hierarchical; root node corresponds to 1 "publication"


Figure from V. Zerkin

Wednesday, November 28, 12
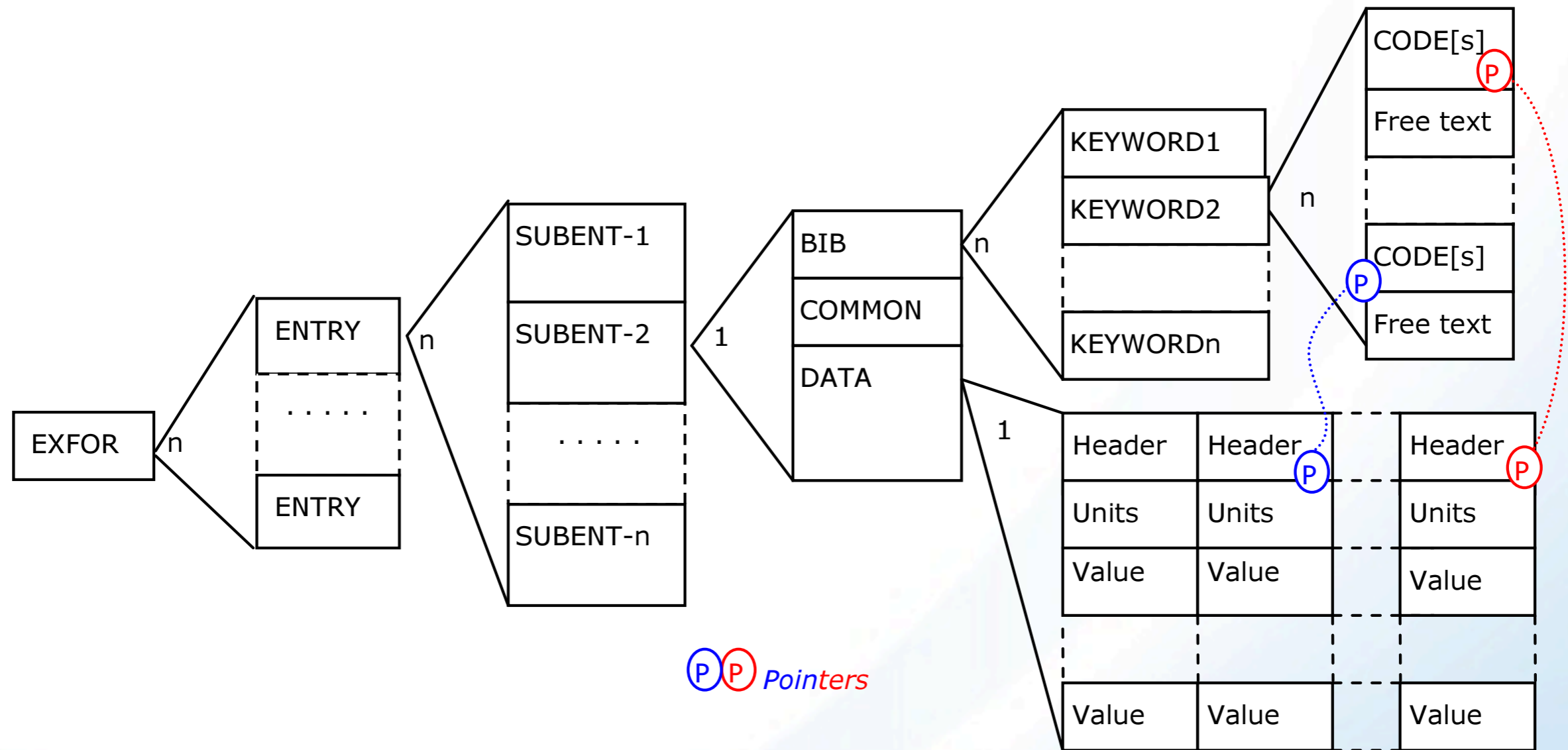
Nevertheless, many "low-level" concepts can and should be reused

- **Bibliographic references**
- **Particles:**
  - Hadrons, elementary (transportable or not)
  - Nuclei
- **Nuclear levels: energy, spin, parity, etc.**
- **Reaction designator**
- **Cross sections**
  - Simple tables
  - Resonance parameters
  - Spectrum weighted
- **Units**
- **xLinks**

- **History/versioning**
- **Documentation**
- **Common, low level, data types:**
  - List
  - Vector
  - Matrix
  - Table
  - Legendre series
- **Higher level constructs**
  - Uncertainties
  - Covariances
- **...**

Wednesday, November 28, 12

But, we cannot guess everything that future users will want

- **Today's models are tomorrow's collision kernels:**
  - FREYA, CoH being integrated into MCNP
  - Enables correlations
  - Needs RIPL....
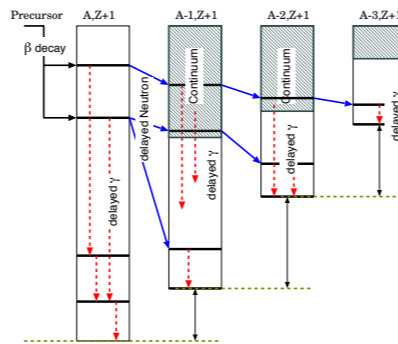
- **Cloud computing:**
  - Assemble new apps from current databases, apps
  - Needs URLs to low level data...
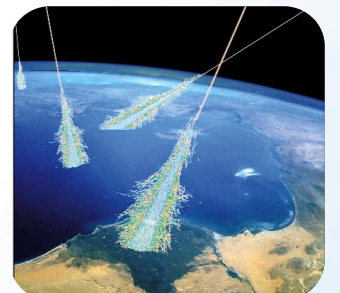 - Relational vs. hierarchical databases

- **Uncertainty Quantification:**
  - Needs support for big data...
  - Ensembles of libraries or covariances or both?

- **But what else will users want?**
  - New particles? neutrinos? muons? heavy-ions? hyperons?
  - All of our databases to (gasp) be in sync with each other?
  - Data mining?
  - Mobile apps?

Optimizing developer happiness optimizes the quality of the products delivered to customers

Brookhaven Science Associates

BROOKHAVEN
NATIONAL LABORATORY

Wednesday, November 28, 12