LEXFOR Entry "Digitization" (A15)

(N. Otsuka, S. Simakov, V. Semkova and S. Dunaeva, 2014-03-25, Memo CP-D/836)

Following the Action 15 from the NRDC 2013 Meeting to one of us (NO), we prepared a draft of a new LEXFOR entry "Digitization". The "Recommendations" and Guideline for Expression of Digitized Data" of the draft are extraction from the summary report of the IAEA Consultant's Meeting on Benchmarking (INDC(NDS)-0629). The first paragraph of the main text is taken from the second paragraph of the LEXFOR entry "Status"- Source of the Data.

We would like to propose addition of the following sentence as the sixth item of "Recommendation":

Quote the digitization uncertainty to two significant digits to avoid an unexpected coincidence in digitized values.

(Background of this addition:
If the rounding is performed based on the digitization uncertainty with one significant digit, it may result in coincidence of two digitized values even if they can be distinguished on the source image by eyes. A typical situation is seen in excitation functions of low-energy charged-particle induced reactions measured by electrostatic accelerators. In such data, the uncertainty due to digitization of the incident energy may dominate over the uncertainties in the projectile energy. (See EXFOR P0132.002 for example). Note that GUM 1995 recommends quoting uncertainties to *at most* two significant digits in general.)

Digitization

When the author's original numerical values have been lost or are not obtainable, data digitized from graphs, if available, should be entered for completeness. Data of this type should be labelled with the status code CURVE.

Example:
```
STATUS      (CURVE) Scanned from Fig. 1 of Yad.Fiz.12(1951)345.
```

Below are some recommendations and guideline agreed in the IAEA Consultant's Meeting on Benchmarking of Digitization Software (Vienna, 12-14 November 2012). See the summary report of the meeting (INDC(NDS)-0629) for more details.

Recommendations

- Ask for numerical data from the authors. Explain to the authors that the original numerical data from the authors are preferable than values digitized from images.

- Avoid adoption of digitized values when the original values are explicitly given in the article (e.g., incident energy, detection angle).

- Enlarge the image as much as possible when the image file is created and digitized. Utilise functions available on the software for this purpose (*e.g.*, "magnifying glass").

- Digitize the beginning and end points of scales with special care to avoid systematic errors.

- Check the values of the beginning and end points (labels on scales) after digitization of the image.

Guideline for Expression of Digitized Data

1. Keep consistency for the number of digits between the digitized values and uncertainties.

   Example:
```
DATA
EN          DATA        DATA-ERR
MEV         MB          MB
 14.1        12.34        2.34
 14.3        12.3         2.3
 14.5         1.234E+01  0.234E+01
 14.6         1.23 E+01  0.23 E+01
 …
ENDDATA
```

2. Use the fixed and floating decimal point expression for the numbers digitized from linear and logarithmic scale, respectively.

   ***Example*:**

   `12.345`   (a value digitized from linear scale)

   `1.2345E+02`   (a value digitized from logarithmic scale)

3. Digitization accuracy may be given in the absolute unit (e.g., `ADEG`) or relative unit (*e.g.*, `PER-CENT`) for the numbers digitized from linear and logarithmic scale, respectively.

   ***Example*:**

   ```
   COMMON
   ANG-ERR-D  ERR-DIG
   ADEG       PER-CENT
    0.12       1.2
   ENDCOMMON
   DATA
   ANG-CM      DATA        DATA-ERR
   ADEG        MB          MB
     5.67       3.456E+02  0.234E+02
    12.31       2.345E+02  0.123E+02
   …
   ENDDATA
   ```

4. Consider rounding of digitized values to integers if values are for atomic numbers, mass numbers etc., and digitized values are close to integers.