

INDC International Nuclear Data Committee

Summary Report of the Consultants' Meeting on **Benchmarking of Digitization Software**

IAEA Headquarters, Vienna, Austria

12 – 14 November 2012

Prepared by

Naohiko Otsuka and Valentina Semkova
IAEA Nuclear Data Section, Vienna, Austria

January 2013

RECOMMENDATIONS

To Software Developers

- R1 (To GDGraph) Make an English version, and make it available for EXFOR compilers.
- R2 (To GSYS) Implement a function to estimate accuracy of digitization.
- R3 (To InpGraph) Resolve too low relative (%) digitization error evaluated by the code.
- R4 Share useful features among developers.
- R5 Consider feedback provided by the meeting participants and compilers.
- R6 Indicate contact details of the developer in the software (*e.g.*, “help”, “about”).
- R7 Implement automatic focusing procedure for symbols, axis and its legend (if it has not been implemented yet).
- R8 Implement feedback function so that compilers can reproduce their digitized values on the images (if it has not been implemented yet).
- R9 Introduce calibration procedure to digitizing programs (if it has not been implemented yet).
- R10 Show digitized scales on the screen independently at the checking stage (if it has not been implemented yet).
- R11 Display a warning message when repetition exists in digitized values of an independent variable.

To EXFOR Compilers

- R12 Follow “Guideline for expression of digitized data in EXFOR” (Appendix C of this report) for digitized data included into EXFOR entries.
- R13 Provide feedback to software developers for further improvement.
- R14 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.
- R15 Do not change the number of digits for values provided by the authors without their permission.
- R16 Avoid adoption of digitized values when the original values are explicitly given in the article (*e.g.*, incident energy, detection angle).

- R17 Consult with NDS if digitization and compilation are really necessary when the image quality is poor, for cases where there have been many experimental data sets in EXFOR.
- R18 Ask for help of other centres when the compiler cannot prepare an image with sufficient quality (*e.g.*, due to restriction to electronic access).
- R19 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”).
- R20 Digitize the beginning and end points of scales with special care to avoid systematic errors.
- R21 Check the values of the beginning and end points (labels on scales) after digitization of the image.
- R22 Investigate the reason of the repetition of a digitized value of an independent variable to avoid human errors such as digitization of the same symbol twice.
- R23 Keep the field for the uncertainty blank when the error-bar is not clearly visible, and explain the situation in free text (with `FLAG` when appropriate).
- R24 Keep text information associated to numerical data (*e.g.*, range of cross sections) in the EXFOR entry as free text when digitized values are adopted in the entry and such information is seen in the article.
- R25 Use InpGraph when deformation of the image is serious.
- R26 Correct defects on objects to be digitized (*e.g.*, by filling symbols or lines) when auto-focusing is applied to a poor image, in order to make auto-focusing more stable.

To NRDC Coordinator

- R27 Add the contact details of software developers to the NRDC software web page.
- R28 Introduce a set of benchmark problems for new EXFOR compilers.
- R29 Introduce a set of benchmark images (*e.g.*, png) for software developers.
- R30 Collect feedback on digitizing software, and list it on the NRDC web page.

GUIDELINE FOR EXPRESSION OF DIGITIZED DATA IN EXFOR

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.