FIRST IASSIST WEST EUROPEAN MEETING
june 27 – 29, 1977

Report of the West European IASSIST Workshop
Held in Copenhagen, June 27-29, 1977

Submitted by

Per Nielsen, Danish Data Archives

Members Present

Europe
Tomasz Bankowski, Zowar Computing Centre, Warsaw, Poland
Flemming Bigom, Danish Data Archives, Copenhagen, Denmark
Merete Watt Boolsen, Danish Data Archives, Copenhagen, Denmark
Ulf Christoffersson, University of Gothenburg, Sweden
J.C. Deheneffe, University of Louvain-la-Neuve, Belgium
Bartlomiej Gasiorowski, Polish Academy of Sciences, Warsaw, Poland
Jørgen Grosbøl, Danish Data Archives, Copenhagen, Denmark
Astrid Bogh Lauritzen, Danish Data Archives, Copenhagen, Denmark
Cees Middendorp, Steinmetzarchief, Amsterdam, the Netherlands
Ekkehard Mochmann, University of Cologne, Federal Republic of Germany
Per Nielsen, Danish Data Archives, Copenhagen, Denmark
Krzysztof Ostrowski, Polish Academy of Sciences, Warsaw, Poland
Karsten Boye Rasmussen, Danish Data Archives, Copenhagen, Denmark

North America
E. M. Avedon, University of Waterloo, Ontario, Canada
Carolyn Geda, University of Michigan, Ann Arbor
Sharon Chapple Henry, Data Clearing House for the Social Sciences, Ottawa, Canada

Agenda

June 27: What do we need to know about each European Archive? How much detailed
          information is really required? Can we design one questionnaire that
          will include all of the required elements and not be a burden to the
          respondent?

June 28: Should archive holdings be standardized with respect to documentation?
          If so, what should be included in these standards? How can we ensure
          that these standards are adhered to and used by primary research per-
          sonnel prior to primary analysis?

June 29: What are the sources of data? Should Archives classify and index these
          data in a standard manner to facilitate data retrieval as well as cross-
          national and international research? If so, how should this be accom-
          plished?

On behalf of the West European IASSIST Secretariat and the DDA, Per Nielsen welcomed the participants and expressed his gratitude that so many "guest invitees" from the East European and North American Regions were present. To bring everybody up-to-date, information was provided on the latest "data-conferences": Carolyn Geda gave information about the IASSIST North American Conference in Florida (February 1977), where 35 data information professionals had attended. Carolyn Geda and Sharon C. Henry then discussed the second North American IASSIST meeting in Toronto (May 1977), in which the number of attendees had doubled compared to the Florida meeting. Working papers and ideas from the above meetings were useful inspiration throughout the Copenhagen meeting. Ekkehard Mochmann reported on the Moscow Conference on Information and Documentation in Social Sciences from which he had just returned; the meeting was an initiative of the Vienna Centre, hosted by INION and supported by UNESCO. The upcoming IASSIST meetings in Chicago (February 1978) and Uppsala (August 1978) were described.

The issue of the first day in Copenhagen was discussed at length. There was agreement that it would be possible to design a Data Organization Registry Form that included all the elements required for a full description of data service organizations. Consequently, it was decided to construct such a Data Organization Registry Form and to recommend this form to IFDO for use; the intended audience for the form was to be the administrators of the various data organizations.

As source material for the Data Organization Registry Form the following documents were used:
1. Questionnaire on data acquisition policies and problems (Marcia Taylor, SSRC Survey Archive, Essex)
2. Questionnaire on data preparation procedures (Eric Tanenbaum, SSRC Survey Archive, Essex)
3. Preliminary list of data elements and subject terms for IASSIST Data Archive Registry in Canada (Lisa Lasko, Institute for Behavioral Research, York University; and Lana Prokop, University of Toronto)
4. Preliminary outline for a "A Guide to Providing Social Science Data Services" (Alice Robbin, University of Wisconsin; and Laine Ruus, University of British Columbia)
5. Request from Alice Robbin and Laine Ruus regarding service documents
6. Input Form (Directory of data bases in the social and behavioral sciences)
7. Data Documentation Form (Data Clearing House for the Social Sciences)

The Group constructed a preliminary questionnaire form consisting of material selected from the above documents; then modified, added, and deleted; and finally ended up with a handwritten version to be computerized before a final review during the Wednesday session.

The working group made an attempt to set up a list of elements describing a data file that would be useful from the user point of view:
(i)   Library information (such as card catalogue)
(ii) Archive information/Study abstract
(iii) Study description (in machine-readable form)
(iv) List of variables (in machine-readable form)
(v)   Codebook (in machine-readable form)
(vi) Classification/index (in machine-readable form)
(vii) Data (not part of documentation, but the basis for the file)
(viii) Special publications (either in print or machine-readable)

Each of the documentation items were subjected to detailed discussions, and some recommendations were agreed upon; a summary is listed below:

(i): Library Information: The North American IASSIST Classification Action Group (coordinated by Sue A. Dodd, University of North Carolina) has set standards for cataloging machine-readable data files. Data organizations and primary research personnel should adopt the library information recommendations of this Action Group.

(ii): Archive information/Study abstract: Each data organization has its own way of publishing its data holdings in an inventory; it seems difficult to make recommendations regarding production of study abstracts. However, it is recommended that data organizations as well as research institutions produce information at the abstract level.

(iii): Study description: It is recommended that the standard study description scheme (developed as a result of a meeting in Copenhagen in June 1974) be tested by several data organizations. As a concrete step in that direction several data holders agreed to test this instrument during the next year:

BASS, Louvain-la-Neuve, Belgium
DDA, Copenhagen, Denmark
ICPSR, Ann Arbor, Michigan, USA
LSDB, Waterloo, Ontario, Canada
Steinmetzarchief, Amsterdam, the Netherlands
Zentralarchiv, Cologne, Federal Republic of Germany
Ulf Christoffersson, University of Gothenburg, Sweden

Five study description questionnaire forms will be disseminated to each of the above holders along with an instruction in the use of the standard study description scheme. Each data holder will send study descriptions (in English language) of five different files to the Zentralarchiv. Preferably, the study descriptions should be submitted on tape; however, the ZA would accept filled-in questionnaire forms from data holders unable to make the study descriptions machine-readable. Attempts will be made to furnish the necessary software for printing the study description in the participating institutions. The outcome of this testing will form the basis for an extension and updating of the instruction manual which is presently being developed and distributed from the DDA. At the end of this process, final recommendations can be made.

(iv): List of variables: It is recommended that the list of variables be available in print as well as in machine-readable form.

(v): Codebook: The "ideal" codebook was outlined by the workshop as consisting of the following 15 items (the first three items referring to file level, the last twelve to be supplied variable by variable where applicable):

(1) Title of study (file and subfile names)
(2) Format of the data-file
(3) Comments at file level (e.g. concerning application of missing data codes; special weighting features; special precautions for use)
(4) Variable identification (number; label; short name; mnemonics)
(5) Variable source reference
(6) Variable location and length
(7) Variable type (alpha; alphameric; numeric; symbolic)
(8) Number of decimal places (scale of measurement)
(9) Source statements/texts/questions/scale descriptions/introductory statements related to the responses

- (10) Code values
- (11) Code descriptions
- (12) Comments: coding and field work (coding and interviewer instructions)
- (13) Variable contingencies (filter; skip; probe; control questions)
- (14) Variable consistency
- (15) Derived variables

(vi): Classification/index: It is recommended that future IASSIST meetings be concerned with this topic. A number of classification schemes (on file and variable level, respectively) are available; however, further testing and elaboration is required before final recommendations can be made.

(vii): Data: Single punch data is preferable for archiving purposes.

(viii): Special publications: It was recognized that many data archives produce publications concerning specific files (examples were examined in hard copy). There was some discussion regarding machine-readable special publications and the future directions that this area of documentation may follow.

The machine-readable preliminary Data Organization Registry Form was reviewed. Questions were added, deleted, modified. This work was finished with agreement on a Data Organization Registry Form to be disseminated to the participants for completion; comments arising from the application of the Form will be reported to the DDA; such comments may make yet another editing process inevitable.

The working group discussed sources of data. This topic had been subject to discussion during the Data Organization Registry Form and Data Documentation sessions of the preceding days; however, the classification/indexing problems regarding data from various sources could not at this meeting be operationalized to a level allowing for recommendations. On the contrary, it is evident that there is some confusion right down to the level of terminology. It would be very useful if the Classification and Process-Produced Data Action Groups of IASSIST could come up with a tentative taxonomy, taking into consideration the increasing number of different data sources relevant to the social science community.

Summary and Prospect
- (1) The Data Organization Registry Form will be disseminated from the DDA to the participants as soon as final editing is completed.
- (2) No later than October 1st participants will complete the Form on behalf of their organizations and return the completed version to the DDA with any comments and suggestions that may arise from the respondent role.
- (3) The Data Organization Registry Form will be recommended to the IFDO for use.
- (4) Concerning the Study Description Scheme participants will receive a manual (Applications & Instructions).
- (5) Participants who have agreed to take part in the testing of the study description will receive five Study Description Forms to be completed in English for five different files in their holdings.
- (6) No later than December 1st, the completed study descriptions will be transferred to the ZA (preferably in machine-readable form) along with a short report on the test experience.