ARCHIVAL OF RADIO SOURCE CATALOGUES: Present Status and Prospects

Heinz Andernach
Instituto de Astrofisica de Canarias
Calle Via Lactea s/n
38200 La Laguna (Tenerife)
Spain

ABSTRACT. Recent proliferation of large radioastronomical source surveys and resulting data products are reviewed, and attention is drawn to their poor archival status. First practical steps are proposed to establish a radio source data base from published source catalogues, proceeding with time from the bigger and recent ones to the smaller and older ones. At later stages the data base could assimilate unpublished data, spectral-line data and also images.

The first and as yet most comprehensive compilation of radio sources was Dixon's Master List of Radio Sources (MSL), first published by Dixon (1970). The most recent update of 1981 (with ~86000 entries) is largely out of date, and many publications (typically those not available in machine-readable form) are not covered. In fact, the MSL was not intended to list all published radio data on any known radio source, and it provides position, flux and reference only. Thus it is of limited use for a systematic retrieval of literature data. In 1977, Westerhout & Jaschek investigated the availability of radio astronomical data in general. The poor response to this inquiry may in part be due to its emphasis on raw and archival data, rather than on published data. Nevertheless, two thirds of the replies to the questionnaire (Westerhout & Jaschek 1989) were in favour of a regular "announcement of new data available" (e.g. in BICDS), but this suggestion remained an idea until now. A recent book on astronomical data (Jaschek 1989), considers radio astronomy a typical area 'inadequately-covered' with data bases, and it concludes that radio astronomers are not interested in creating a data base. Here we intend to overcome this and suggest practical solutions. Some recent attempts to build up source data bases (e.g. CERCO80-84 (Gopal-Krishna et al. 1985) or PKSCAT90 Otrupcek & Wright 1991) were limited in either the covered epoch of publication or in the covered region of the sky. The author is preparing a bibliography to keep track of radio source data, which are scattered over hundreds of publications. The total number of published radio source measurements amounts to >500,000 in mid-1991. The rule "many papers with few data, few papers with many data" (Jaschek 1989) holds perfectly: a representative list of the 28 largest source catalogs comprises ~350,000 entries. It reveals two especially alarming features: None of the 28 catalogs are available from the established data centers! The amount of data published since 1990 until mid-1991 exceeds that of the entire eighties! Considering ongoing and planned surveys we can anticipate an average growth rate of ~200,000/yr until the year 2000, by which we expect the number of measured flux densities to be five times the present one.

J. Bergeron (ed.), Highlights of Astronomy, Vol. 9, 731–733.

Attention to this problem was drawn by the author in an e-mail campaign since late 1989 (see Andernach 1990 for the first results). Apart from a disappointing lack of response from the major radioastronomical institutes, the reaction of individuals was dominated by both affirmative interest and reluctance, proving the awareness of a high demand of manpower and expertise required to create a comprehensive data base. Positive aspects are however, that - a large fraction of existing data is contained in few machine-readable files. Some 16 of the above 28 major source lists (with 218,600 entries) and another dozen minor ones were obtained from the authors. However, other lists of appreciable size turned out to be either unavailable or lost, resp. erased. - modern page scanners could be used to transfer the missing source lists into machine-readable form. - expertise from other ongoing projects (SIMBAD, NED, DIRA, PKSCAT90, PGC, RC3 etc.) may be consulted before setting up appropriate software for a radio source data base. There is no need to reinvent the wheel!

The author's inquiry has also shown that in none of the many existing data centers any efforts are being invested towards a radio source data base. However, during two informal meetings, arranged by the author in July 1991 (at MRAO Cambridge (UK) and during the General Assembly in Buenos Aires), agreement was reached on the following practical steps: - The source lists collected so far, will be gradually incorporated within two existing data bases (EINLINE at CfA and DIRA at Bologna). By advertising this service, it is hoped that more authors will contribute their published data. The author volunteers as a provisional 'clearing house' for such data and for any relevant information on "what is available where". A working group of IAU Comm. 40, entitled 'Archival and Databases for Radio Astronomy' and chaired by the author, was created during the General Assembly. The WG should discuss the scope of the data base project (its basic structure: which parameters to include, priority of data inclusion, growth potential, etc.; hardware needed: computers & page scanner, typing aids). On the long term, and for the sake of reliability, the source data base should be installed and maintained in a center of radioastronomical expertise. The WG should try to help define such a place and consider a proposal for financial support to a suitable funding organization. - Every attempt will be made to maintain original literature data within the data base. Any evidence for data being superseded or in error, should be documented properly (i.e. transparent to the user), to avoid a loss or change of original information. Any systematic exclusion from the data base of radio sources by type (as e.g. galactic sources in NED (Helou et al. 1990)) should be left to the users, as it is a time-consuming and often a subjective task. - To avoid discrimation of developing countries, the data should be accessible free of charge to all interested astronomers, as is e.g. practised by NED. - At later stages the data base could assimilate unpublished data (if they are sufficiently well documented), and extended to cover spectral-line data and images (from individual sources up to the entire sky). Explicit interest for the latter transpired at the XXIst IAU General Assembly.

A few remarks to put the envisaged radio source data base in perspective: Heterogeneous compilations like Dixon's Masterlist are sometimes used for inadequate statistical analyses with questionable results, and the proposed data base obviously bears the same danger. However, the data base is meant to be a tool for retrieval of literature data, and the users themselves are responsible for a careful avaliation (or correction) of the retrieved data. - Unlike a photographic plate or a CCD frame in optical astronomy, raw data in radio astronomy are generally useful only for the (typically few) astronomers familiar with the reduction software specific to the observational method used. For the most notable exceptions (VLA and WSRT data reduced with AIPS) the raw data are being adequately archived by NRAO and at Westerbork. While archiving of raw and unprocessed data is

considered indispensable, it is not the aim to include them here, except maybe for hints of their existence. Variability is one, but not the strongest argument for the present proposal. The main argument is to offer a quick answer to the common question: which radio (and later other) data are known for a given radio source, specified by either position or name. The data should be stored with all necessary technical information on the telescopes used. However, in radio astronomy as in any part of astronomy, this item is limited by the publications themselves, which not infrequently omit a statement of whether a listed flux density is peak or integrated, whether a displayed map has been primary beam corrected or not, up to the missing specification of the contour levels.

In conclusion, the last sentence of the paper with the largest individual source list published so far (Becker et al. 1991) might serve as a proof that radio astronomers are in fact feeling the need for 'their' data base: "The challenge in the future will be at least as much in assimilating as in collecting data."

REFERENCES

Andernach, H., 1990, Bull.Inf.CDS 38, 69
Becker, R.H., White, R.L., Edwards, A.L., 1991, ApJS 75, 1
Dixon, R.S., 1970, ApJS 20, 1
Gopal-Krishna et al. 1985, 'Compendium of Extragal. Radio Continuum Observations', magnetic tape version advertised on IAU XIX, New Delhi
Helou, G., Madore, B.F., Bicay, M.D., Schmitz, M., Liang, J., 1990, 'Windows On Galaxies' (eds. G.Fabbiano et al.), Kluwer Acad.Publ., p. 109
Jaschek, C., 1989, 'Data in Astronomy', Cambridge Univ. Press
Otrupcek, R.E., Wright, A.E., 1991, Proc ASA 9 (in press)
Westerhout, G., Jaschek, C.: 1978b, Bull.Inf.CDS 15, 99