# REMARKS ABOUT THE CATALOGUING OF OPEN CLUSTER DATA

Gösta Lyngå
Lund Observatory
Box 1107
S-221 04 Lund
Sweden

During the last three years I have collected and examined data for open clusters. I have chosen to make a computer-based catalogue. Fig. 1 shows the structure of this work. The data reside on a disc at the Lund Observatory. As original input I have used the data banks available in 1980. Later information is added through frequent updates. There are two modes of retrieving the data: by listing and by direct access.

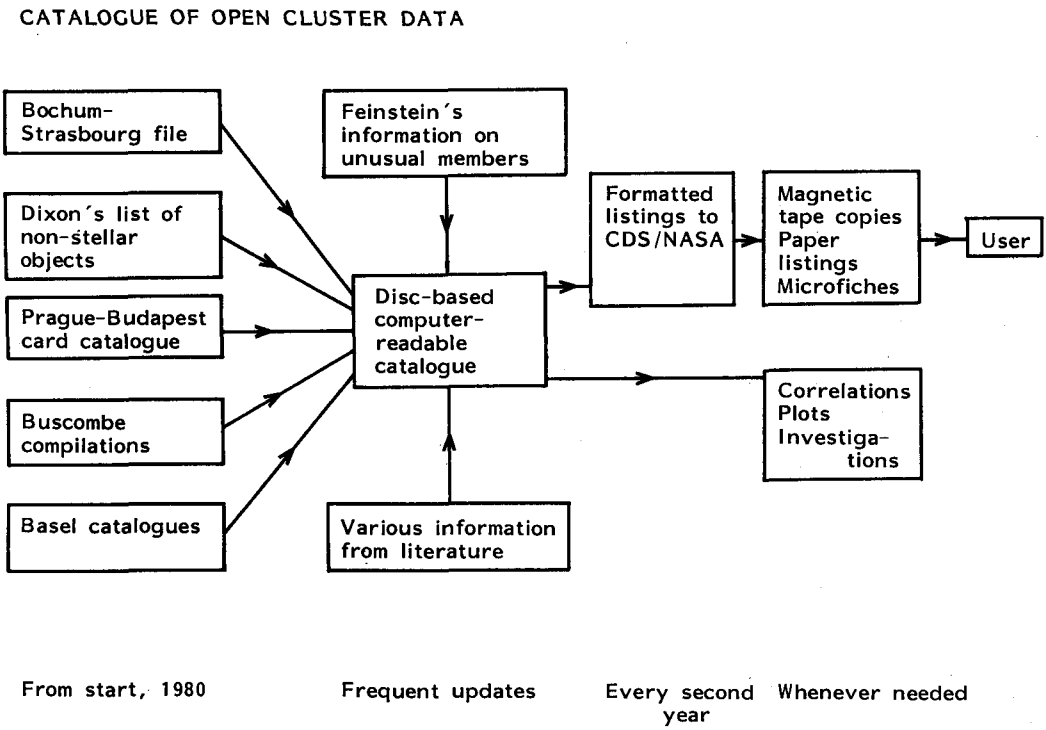


Fig. 1. Cluster catalogue lay-out.



*C. Jaschek and W. Heintz (eds.), Automated Data Retrieval in Astronomy, 289–291.*

## I. LISTING OPTIONS

One way of retrieving the data is to make a listing and I do so every second year in order to distribute the data as it is known by the computer base at that stage. The distribution goes via CDS and NASA and can be made available to the user in different ways. Below are listed the files that are stored on the magnetic tape. Files 1 - 5 are in ASCII code with a block length of 7920 characters corresponding to a page with 60 lines, each containing 132 characters. File 6 is in HP format and is a direct copy of my own disc at the time of submitting the catalogue. This file is not aimed for printing but rather for computer-controlled investigations of cluster data.

TAPE FILES

| | | |
|---|---|---|
| FILE 1 | INTRODUCTION DESCRIPTION | 0.02 Mbyte |
| FILE 2 | FORMATTED CATALOGUE | 0.95 |
| FILE 3 | REFERENCES | 0.05 |
| FILE 4 | ALIAS NAMES | 0.21 |
| FILE 5 | TRACED TABLES | 0.35 |
| FILE 6 | CATALOGUE IN HP FORMAT | 1.21 |

## II. COMPUTER CONTROLLABLE FILE

So far I have only described what could have been a printed catalogue, albeit one for which new editions can be frequent. If that were all that could be achieved with a computer, it would not be worth the effort. The far more important aspect is that it is possible to interact with the catalogue, changing parameters, deleting entries and adding new information. Thus, while a printed catalogue is ageing from the time of its submission, a computer-based catalogue is growing and changing continuously. Other advantages are the flexibility of list structures (Davis, 1977), the large storage capacity and the mobility of the data bank.

It has been my aim to create the sort of structure for the catalogue of open cluster data. Below I have summed up the advantages and the problems encountered so far. The advantages are obvious and I hope to show their value elsewhere, in connection with studies of galactic structure.

COMPUTER-BASED VERSUS PRINTED CATALOGUE

| ADVANTAGES | PROBLEMS |
|---|---|
| Ease of revision | Inhomogeneity |
| Direct access | Notation of imprecision |
| Error removal | Inflexible formatting |
| Flexible ordering | Program errors |
| Large storage | Unfamiliarity |
| Mobility | |

https://doi.org/10.1017/S0252921100083135 Published online by Cambridge University Press

III. PROBLEMS

In the present context the problems are of interest. The most fundamental difficulty with all large collections of data values is the inhomogeneity. A data value may be different depending on telescope, exposure time or other observing conditions. It may also refer to a different interpretation of observed object. In a printed catalogue such details may be remarked on. Often such notes are, however, not made and then the printed catalogue has no advantage over the computer-based one. In the Prague-Budapest catalogue all published values are given together with their references. This may sometimes be desirable but it can be difficult for an investigator who is not familiar with the astronomical literature. I have for many parameters tried to select one single value for each cluster. Such a selection is necessary for applications but it may still suffer from a certain degree of inhomogeneity.

Connected with this problem is the one of assigning imprecision. An example: the radial velocity of a cluster is listed as one value, expressed in km/s. It may be obtained as a mean value of twenty accurately measured radial velocities or refer to a rough measurement of one possible member star. Yet, what we enter into the data base is one value per cluster. In this as in many other cases there ought to be some standard way of denoting quality of data. The printed catalogue can use a number of methods like colons, italics, underlining or variable number of decimals to give the reader a guide to his selection of data. In the computer-based catalogue such a function must be filled by a separate parameter giving weights or assigned values of imprecision.

Let me in this connection also make a comment about the limitations of format imposed by computers. If the colour excess of a cluster is $0^m.12$ I want to list that value. Another cluster has the less well known colour excess of $0^m.1$; this will appear as $0^m.10$ if the format (F4.2) is used. To beat the system I then decide to use free field format and the colour excesses are given as $0^m.12$ and $0^m.1$. All very well, but what about the cluster for which accurate photometry gave $0^m.10$? That also will be listed as $0^m.1$.

IV. CONCLUSIONS

Let me sum up by pointing out that computers have given us entirely new ways of comparing data and of drawing conclusions. However, to know the extent of possible bias and unreliability we must be in full control of the procedures. At the present, there appears to be no more efficient way to interrelate the large amounts of available observations than to use computer-based data banks. Thus it is imperative to develop good standards for critical inspection of the data.

References

Davis, M.S. 1977: In Jaschek, Wilkins: Compilation, Critical Evaluation and Distribution of Stellar Data. IAU Colloquium No. 35, p. 3.

Lyngå, G. 1981: Astronomical Data Center Bulletin, Vol. 2. In press.