PART IV

THE DISTRIBUTION OF DATA

https://doi.org/10.1017/S025292110005301X Published online by Cambridge University Press

DATA DISTRIBUTION

B. Hauck

Institut d'Astronomie de l'Université de Lausanne
et
Observatoire de Genève, Switzerland

I. Introduction

Before answering, or trying to answer, the question "how to distribute astronomical and astrophysical data", it is necessary to reply briefly to the question "why distribute these data?" This question is not so trivial as it first appears and it even hides the basic problem.

Data acquisition in itself is a single or manifold technique, often very delicate, constituting in itself a whole field of research. The interpretation of data follows it and will not necessarily be made by the same people. Very often only one possible aspect will be examined. It is consequently absolutely necessary that the data should be disseminated as widely as possible in order that they are used to the utmost. If it is necessary to assure the dissemination of data, it is no less necessary to ascertain their conservation. I shall give here only one example: the observations which have permitted to obtain the light curve of the quasar 3C273.

The necessity to distribute and to preserve the data being established, we can now examine the problems related to their distribution.

If a session of this colloquium is devoted to data distribution, it is because the traditional method of printing a catalogue or a list, and forwarding it to other observatories by some way or another is certainly no longer well suited to the present situation. On the one hand because the number of new data to be published increases at an extraordinarily high rate and, on the other

C. Jaschek and G. A. Wilkins (eds.), Compilation, Critical Evaluation, and Distribution of Stellar Data. 155-159.

hand, because there exist today other means. It is also necessary to take into account the fact that it is often important to have new data available very quickly and that some data have a short lifetime.

II. Increase of Data

C. Jaschek (1968, 1973) has already demonstrated the very large increase, or inflation, of data in various fields of stellar astronomy and it is not necessary for me to give once more this information. However, I shall give you briefly some statistics concerning the photometric data. These are given in table 1. The numbers given in this table appear to me to be a good demonstration of the necessity to consider new methods for data distribution.

III. Methods of Distribution

In this section I shall examine essentially the question of supports for the information. It is clear that the classical way - that of printing catalogues and lists of data - remains, and will remain, a very useful and valid method. Some journals play an important part in this direction, in particular the Astrophysical Journal and Astronomy and Astrophysics, in editing Supplement Series intended for the dissemination of the data. The increase of the latter brings about an increase in the number of published pages. The Supplement Series of Astronomy and Astrophysics contained 1075 pages in 1970 and 2979 in 1974! This leads also in many cases to an increased delay of publication. It is still necessary to add some catalogues edited directly by the observatories.

We may now ask ourselves whether this support is the only one to maintain. My answer is no. It is indeed very useful to have available on a shelf in one's own office catalogues which are easily consulted and this fact must be taken into consideration. However it is also necessary to bear in mind with regard to this kind of distribution the fact that an up-to-date version is very difficult to obtain and that the problem of the inflation of printed pages is more acute. Furthermore, it is not possible to use directly the data in this form with a computer. This is certainly the greatest weakness of data distribution only in printed form.

Today it seems to me that it is essential that the basic support of catalogues should be magnetic tapes. It will then be possible to distribute them to people interested in using them in some computations. It is certainly easy to object that not everyone can work with magnetic tapes, but I wish to insist on the fact that if I consider the magnetic tape as the best suitable basic

support for the information, I think nevertheless that it is not the only means of distribution. If you have a catalogue on magnetic tape, then you have at your disposal many various ways of distribution, such as

(a) a copy of the master tape
(b) a listing (but this process is expensive)
(c) offset copies
(d) microfiches

In the latter case, you certainly know of the existence of machines which produce microfiches directly from a tape.

Thus, a catalogue on magnetic tape cannot only be used with a computer, but can also be distributed in numerous other ways. To the use of the various supports mentioned, I wish to add the possibility to use a catalogue on magnetic tape in the framework of a system of remote computing, the user being linked to a central computer by only a console or through its own computer. Such linkage of communications can exist either inside an observatory or on a national (or international) scale. The example of the INAG Computer network in France (Jung 1974) is a very good one.

IV. A Possible Solution for Distribution

Let us now examine a possible solution for distribution which will take simultaneously into account the wishes of some people (quick access, if possible catalogue on library shelves) and of others (possibility to treat the information directly on a computer). It is also necessary to take into consideration the problems of the editors and to avoid the congestion of journals!

A catalogue must firstly be put by the authors on magnetic tape, then it will be sent to the editor (with copies of listings) of a journal publishing data. It is important also that a catalogue be submitted to a referee. In some cases, the catalogue will be reproduced by offset, but in most cases a microfiche would be made and attached to the issue containing a short description of the catalogue. Thus, the existence of the catalogue would be known by all interested people. Moreover, it would be possible to give it an accurate reference in subsequent papers. This is a very important point for the authors. The catalogue would also have to go through the filter of a referee system, which is important if we wish to avoid a proliferation of bad catalogues. As to the magnetic tape, it would be sent after the microfiche process by the editor to a data centre. The centre is responsible for the distribution of copies and the preservation of the master copy. It is important that the astronomical community is guaranteed the

distribution and storage of the tapes, and only the collaboration between a journal and a data centre can assure their distribution and storage.

At the present time, an intermediate solution is proposed by the Supplement Series of Astronomy and Astrophysics. An author can submit a short description of his catalogue. Only this description and one page of the listing are printed; but the description and catalogue (in the form of a complete listing) are examined by a referee. The tape is deposited at the Stellar Data Centre at Strasbourg. This centre will store the master copy and distribute copies, listings or microfiches.

V. Data Centre and Data Distribution

During this paper we have seen that an important part is devoted to data centres. We can now discuss the links between the various centres. It seems to me out of the question to envisage only one centre involved with all astronomical and astrophysical data. Such a centre would be rather overloaded and I doubt its efficiency. Now a data centre is useful only if it is efficient, in particular by distributing data very rapidly. I return to a solution which I proposed some time ago (Hauck 1974). A centre should be specialized in a very well-defined field (Stellar Data Centre, for example). It would have the responsibility of the keeping up-to-date, the distribution and the storage of data connected with its activity. In order to improve the diffusion in defined geographic areas, some sub-centres would be in charge of distribution in their own area. They should also be remotely accessible. Their files would be regularly brought up-to-date by the main centre responsible for the field concerned. These sub-centres could act simultaneously as sub-centres for many fields.

VI. Conclusion

The present possibilities of diffusion and treatment of data certainly give us a valuable advantage compared with our predecessors. It is possible to obtain up-to-date data regularly which are easily accessible, on supports allowing various uses. But it is nevertheless necessary to have the willingness to use what is at our disposal and to make sure that a coherent policy exists or is applied between the various data centres (and sub-centres) together with close collaboration with specialized journals.

Table 1

I. *UBV* SYSTEM	
Blanco, V.M. *et al.* (1970)	24582 entries
Mermilliod, J.-C. and Nicolet, B. (1976)	72943 entries
II. *uvbyβ* SYSTEM	
Strömgren, B. and Perry, C.L. (1965)	1217 stars
Lindemann, E. and Hauck, B. (1973)	7603 stars
Hauck, B. and Mermilliod, M. (1975)	9407 stars
Hauck, B. and Mermilliod, M. (1976)	14589 stars
III. GENEVA SYSTEM	
Rufener, F. *et al.* (1964)	342 stars
Rufener, F. *et al.* (1966)	686 stars
Rufener, F. (1971)	1406 stars
Rufener, F. (1976)	4670 stars

References

Blanco, V.M., Demers, S., Douglass, G.G. and FitzGerald, M.P.: 1970, *Publ. U.S. Naval Obs.*, 2nd series, 21.
Hauck, B.: 1974, *Bulletin CDS*, Strasbourg, no. 6, 1.
Hauck, B. and Mermilliod, M.: 1975, *Astron. Astrophys. Suppl.* 22, 235.
Jaschek, C.: 1968, *Publ. Astron. Soc. Pacific* 80, 654.
Jaschek, C.: 1973, in Ch. Fehrenbach and B.E. Westerlund (eds.), *IAU Symp. 50*, p. 275.
Jung, J.: 1974, *Bulletin CDS*, Strasbourg, no. 7, 2.
Lindemann, E. and Hauck, B.: 1973, *Astron. Astrophys. Suppl.* 11, 119.
Mermilliod, J.-C. and Nicolet, B.: 1976, in preparation.
Rufener, F., Hauck, B., Goy, G., Peytremann, E. and Golay, M.: 1964, *Publ. Obs. Genève*, série A, no. 66.
Rufener, F., Hauck, B., Goy, G., Peytremann, E. and Maeder, A.: 1966, *J. Obs.* 49, 417.
Rufener, F.: 1971, *Astron. Astrophys.* 3, 181.
Rufener, F.: 1976, *Astron. Astrophys.* (in press).
Strömgren, B. and Perry, C.L.: 1965, Institute for Advanced Study, Princeton, no. J (second version, unpublished).