

S.I.M.B.A.D. STORY:
A DESCRIPTION OF THE DATA BASE OF THE STRASBOURG
STELLAR DATA CENTER

D. EGRET
Observatoire de Strasbourg

Abstract : The aim of this paper is to give the information necessary for those who want to make extensive use of the CDS Data Base: its history, content, sources, statistics, existing editions, bibliography.

1. INTRODUCTION

The Data Base of the Stellar Data Center (Centre de Données Stellaires; CDS) at Strasbourg is organized as a set of interconnected files, accessible on-line on permanent disks, under the generic name of SIMBAD (Set of Identifications, Measurements, and Bibliography for Astronomical Data) on the Univac 1110 computer of the Centre de Calcul de Strasbourg-Cronenbourg.

The major problem which appears immediately when merging star catalogs is that each set of data may use a different star designation system, so that one given object may have ten or more different names (see for instance Jaschek et al., 1980). For this reason, the central file of the Data Base consists of the Catalog of Stellar Identifications (CSI) : this is a dictionary of stellar designations containing identifications and fundamental data for about 440 000 stars. The CSI was created by J. Jung and his collaborators on the IBM 360/65 computer of the Institut National d'Astronomie et de Géophysique (see Jung and Bischoff, 1971; Ochsenein et al., 1977), as a first step toward the collection of stellar data. A microfiche edition of the CSI and some connected files was prepared in 1979 (Ochsenein et al., 1981).

A number of fundamental changes have been introduced since this last edition, when the Data Base was transferred from one site and computer (Meudon; IBM 360/65) to another (Strasbourg-Cronenbourg; Univac 1110). The aim of this paper is to give all the information necessary for those who want to make extensive use of the CDS Data Base. I will describe the present organization of SIMBAD as well as the CSI as it is distributed on microfiche or magnetic tape.

2. CONTENT OF THE DATA BASE

The CSI resulted from a merging of the following basic catalogs: the SAO Catalog, the HD, the AGK2/3, the Cape Photographic Catalogue (-30° to -40° and -52° to -90°) supplemented by the Cape Zone Catalogue (-40° to -52°), the Yale Zone Cata-

logues (+50° to +60° and -30° to +30°), the Cape Catalogue of Faint Stars (-40° to -52°), and the General Catalogue (GC).

In a second step, a number of other catalogs (e.g. the IDS, N30, GCRV, UBV, etc.; see Table 1) have been linked to the CSI: this means that for each star in the CSI, a flag indicates whether the star figures in a given "linked" catalog, although generally not all the stars of such catalogs could be included (because of a lack of precise coordinates, ambiguous designation, etc.). The list of connected catalogs has steadily increased: at the time of the microfiche edition (Ochsenbein et al., 1981) about 36 catalogs were connected, and the CSI contained 434 946 entries.

Table 1
Catalogs connected to the CSI (1982). The complete references to the catalogs are given in the paper by Ochsenbein et al. (1981)

CATALOG	DATA INCLUDED IN THE DATA BASE	NUMBER OF OBJECTS	CONNECTION RATIO
A) Identifications and positions			
DM	Durchmusterungen (BD, CD, CPD)	no	410 000 ~40%
*	Bayer and Flamsteed designations	no	
CPC	Cape Photographic Catalog (+Cape Zones)	(positions) 87 000	all
CF	Cape Catalog of Faint Stars (1939)	" 20 000	all
HERZ	Sternkatalog für die Zone -6 bis -10° (1907)	" 10 200	all
NGC, IC	Designations in open clusters	no	
*V	General Cat. of Variable Stars (Kukarkin et al.)	no 3 000	15%
ADS/IDS	Catalogs of Visual Double Stars	no 40 000	50% (ADS:70%)
B) Astrometric catalogs			
S.D	SAD (1966)	positions + μ	259 000 all
AG	AGK2/3 (Lacroute and Valbousquet, 1971)	μ	180 000 all
YZ	Yale Zones (+50 to +60° and -30 to +30°)	(positions) 137 000	all
FK4	FK4 and FK4 suppl.	no 3 500	all
N30	N30	positions + μ	5 300 all
GC	General Catalogue (B. Boss, 1936)	μ	33 000 all
PLX	General Cat. of Trig. Parallaxes (Jenkins)&Suppl	μ 6 300	all
C) Photometric catalogs			
UBV	UBV Photometry (Blanco, 1968; Mermilliod, 1978)	UBV 40 000	70% (+)
GEN	Geneva Photometry (Rufener, 1976 ed.)	UBVB ₁ B ₂ V ₁ G 4 500	95% (+)
UVBY	uvby δ Photometry (Hauck and Mermilliod, 1975)	uvby δ 11 000	95% (+)
CFL	Celelescope Catalog of Ultraviolet Photometry	U ₁ U ₂ U ₃ U ₄ 5 700	95%
IRC	2 μ Sky Survey (1969)	K, I 4 100	70%
JP11	UBVRIJKLMNH Photometry (Johnson)	UBVRIJKLMNH 4 500	95%
D) Spectroscopic catalogs			
HD/HDE	Henry Draper Catalog and extension	Sp. 260 000	all (HD)
MKJ	MK Classifications (La Plata & extension 1978)	MK class. 31 000	~ all (+)
MSS	Michigan Spectral Survey for the HD Stars (1 & 2)	" 66 000	all (+)
EM*	Cat. of emission stars (Mackerling, 1970)	no 4 000	40%
HGAM	H γ Measures (Petrie et al., 1973)	W(H γ) 1 200	all
GCRV	Catalogs of Radial Velocities: Wilson, 1953; Evans, 1967; Abt and Biggs, 1972; Barbier, 1975.	RV 26 600	90%
ROT	Catalogs of $V \sin i$ (Bernacca & Uesugi, Fukuda)	$V \sin i$ 4 400	95% (+)
LS	Luminous Stars in the Milky Way (North & South)	no 9 000	75%
FE/H	Abundance Determinations (Morel et al., 1975)	no 540	60% (+)
SB	Spectroscopic Binaries (Batten, 1967)	no 730	75%
E) Combined data			
HR	Catalog of Bright Stars (Hoffleit, 1964)	no 9 100	all
GJ	Catalog of Nearby Stars (Gliese, 1969)	r, M γ , U, V, W 1 890	all

(+) New editions are now available and will be integrated in the future.

In addition to the CSI another major file -the Bibliographical Star Index (BSI)- contributes to the Data Base (see 4.2, below). In early 1983, the BSI contained approximately 100 000 entries, 80% of them being already included in the CSI. This means that for the remaining 20 000 objects only bibliographical data are available. An effort is being made to find cross-identifications or precise positions for these stars in order to include them in the CSI. In the following we will consider generally only those stars for which coordinates are available in the Data Base (i.e. only the stars included in the CSI).

3. THE FUNDAMENTAL DATA

For each star the Data Base gives the DM number (or key identification), coordinates, magnitudes, spectral type, proper motions, and identifications (or running numbers) in the linked catalogs, as discussed in the following paragraphs.

3.1 The key identification : the key number (necessary in the first version of the CSI, which was stored on indexed sequential files) is the DM number, in three parts: declination zone (in degrees), running number in the zone, and code. This code (two digits) is used to indicate which Durchmusterung is used (Bonner, Córdoba, or Cape Photographic Durchmusterung) and, simultaneously, to discriminate stars having the same DM number (for example in a multiple system). This code is 0 for a BD number, 20 for a CoD number, 40 for CPD, when the star is a simple star. It is 1, 2, 3, etc. (giving respectively 21, 22 or 41, 42, etc.) to distinguish the first, second, third, etc. component in a multiple system (see 3.6, below).

Example : CoD -42 1587 (component A of double) : key number -42 1587 21
key number of component B : -42 1587 22

In the zones where two DM overlap, both numbers are kept (when known), because many authors do not follow the HD convention (BD north of -23° , CoD between -23° and -52° , CPD south of -52°). Both numbers can be used as an entry to the Data Base.

Example : CoD -42 1587 = CPD -42 513
CSI key number : -42 1587 21

Not all the stars in the CSI figure in one of the Durchmusterungen: for approximately 5% of the stars presently included it was therefore necessary to create a *CSI number* (used as the key number in the files), or "pseudo-DM", also consisting of three parts: declination zone, number in the zone, and code. The declination zone (in degrees) is computed for the equinox 1950 (while it is 1855 for the BD and 1875 for the CoD and the CPD); the number in the zone is, in this case, *negative* in order to distinguish it clearly from a real DM number, and has 5 digits reproducing the right ascension of the star (at equinox 1950.0) to a tenth of a minute.

Example : BSD $+40^\circ$ 2004; no BD number
coordinates (1950) : $20^h 51.8$ $+45^\circ 46$
CSI number: +45 -20518 0.

Here again, the code discriminates between stars with the same right ascension (to 1/10 minute) and in the same declination zone; it is a running number assigned to the star when it is incorporated into the CSI: 0, if the star is alone in the given range of coordinates; 1, 2, 3, etc. otherwise (in the case of multiple systems, clusters, etc.).

Example : in NGC 2451 (list from Williams, 1967)

star no. 8	$7^h 43.0$	$-38^\circ 4$	CSI: -38 -7430 1
star no. 37	$7^h 43.0$	$-38^\circ 2$	CSI: -38 -7430 2

In the present organization of the data files, on the Strasbourg site, such a key number is no longer necessary from a processing point of view. Assigning a designation to all stars in a common system is nevertheless one of the only ways to *treat redundancies* safely. In any case the "pseudo-DM" number is to be kept for *internal use* on SIMBAD, in order to avoid increasing the existing confusion in the field of star designation.

3.2 The coordinates and proper motions: the coordinates (equatorial coordinates at equinox 1950.0) are from the basic catalogs listed above (section 2), the majority from the SAO catalog. For about 10% of the stars, only HD coordinates are available (given to the arc-minute). In the microfiche edition the precision was not better than 1/12 sec in α and 1.3 arcsec in δ because the coordinates were originally stored in a 40-bit string. In the present version of the Data Base the coordinates from the SAO Catalog are restored, with their full precision. The nominal precision (0.01", 0.1", 1", 0.1', 1') is flagged.

Presently the proper motions are from: 1) the SAO Catalog; 2) the AGK3 (Lacroute and Valbousquet, 1970). The mean errors are also given. When one is interested more specifically in astrometric data, it is possible to check in the list of identifications (see below: 3.5) to see whether the star figures in a connected astrometric catalog (such as the FK4, N30, or AGK3) and, if so, to retrieve the corresponding data from the disk file or the tape (see 4.1). Some 360 000 stars from the CSI figure in an astrometric catalog (see Ochsenbein et al., 1981, Table II). For these stars it is also possible to decide an order of priority (e.g. : FK4, N30, AGK3, GC, etc.) and to retrieve the corresponding published proper motions, which may, then, differ from the "fundamental" proper motions recorded in the main part of the Data Base.

3.3 The B and V magnitudes: the magnitudes come from the UBV photometry and a number of other sources; Ochsenbein (1974a) homogenized them with respect to the UBV system; the sources and estimated mean errors (generally ± 0.2 mag.) are examined in the same paper. The duplicity and the variability are flagged (when known, on the basis of the linked catalogs) and the nominal precision is given. At present, 7% of the stars in the CSI have no V magnitude, and 9% have no B. Nearly 8000 stars (2%) have neither a V nor a B magnitude known.

Some features of the distribution of the CSI stars in V magnitudes have been presented by Ochsenbein (1976); the histogram of visual magnitudes is reproduced as Fig.1

Figure 1 :

Distribution of CSI stars according to visual magnitudes (from Ochsenbein, 1976)

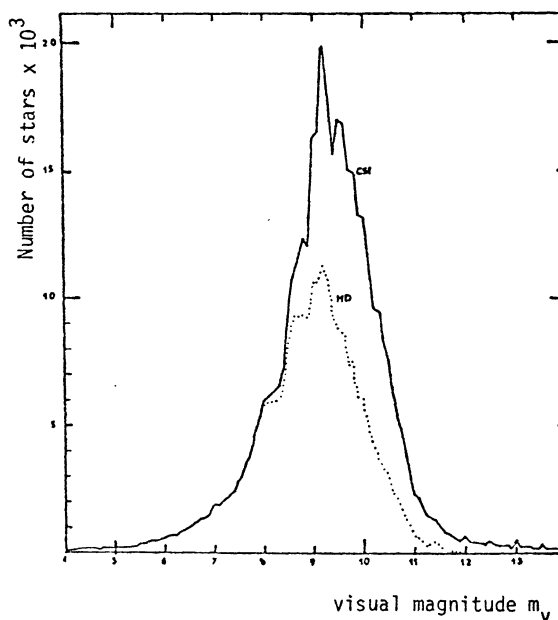


Table 2

CSI statistics on visual magnitudes
(October 1980; additional entries
in multiple systems are omitted)

mV _i	Number of stars with mV ≤ mV _i	Number of stars with mV _{i-1} < mV ≤ mV _i
6.0	5 458	-
6.5	9 437	3 979
7.0	16 542	7 105
7.5	27 766	11 224
8.0	50 110	22 344
8.5	85 659	35 549
9.0	148 582	62 923
9.5	235 763	87 181
10.0	309 709	73 946
10.5	357 123	47 414
11.0	379 910	22 787
11.5	388 831	8 921
12.0	392 909	4 078
12.5	395 290	2 381
13.0	397 285	1 995
14.0	399 713	2 428
>		1 167
TOTAL	400 880	
No mV known :	33 143	

Table 3

CSI statistics on blue magnitudes
(October 1980; additional entries
in multiple systems are omitted)

mB _i	Number of stars with mB ≤ mB _i	Number of stars with mB _{i-1} < mB ≤ mB _i
6.0	3 328	-
6.5	5 670	2 342
7.0	9 240	3 570
7.5	15 774	6 534
8.0	26 726	10 952
8.5	45 778	19 052
9.0	75 627	29 849
9.5	122 839	47 212
10.0	196 097	73 258
10.5	267 737	71 640
11.0	330 527	62 790
11.5	362 919	32 392
12.0	379 970	17 051
12.5	386 485	6 515
13.0	389 254	2 769
14.0	391 590	2 336
>		1 216
TOTAL	392 806	
No mB known :	41 217	

and more detailed information is given in Tables 2 (V magnitudes) and 3 (B magnitudes). When more specific magnitudes or colors are needed, one should look for UBV photometry, which can be retrieved automatically for some 40 000 stars. Other photometric catalogs are also connected to the CSI (see Table 1), up to a total of 48 000 stars (see Table II of Ochsenbein et al., 1981).

3.4 The spectral type: the MK classification is recorded for some 95 000 stars from the volumes I and II of the Michigan Spectral Survey and the Catalog of Selected MK Types (Jaschek, 1978). In the absence of MK classification, the HD spectral type is recorded. No spectral information is available for 8% of the CSI stars. Here again the cross-identifications can be used in order to retrieve other published MK classifications, in addition to the "fundamental" one.

Some features of the distribution of the CSI stars versus HD spectral type were presented by Ochsenbein (1976) and the histogram is reproduced here as Fig. 2. The detailed statistics of B magnitudes versus spectral type were published by Nagy et al. (1981).

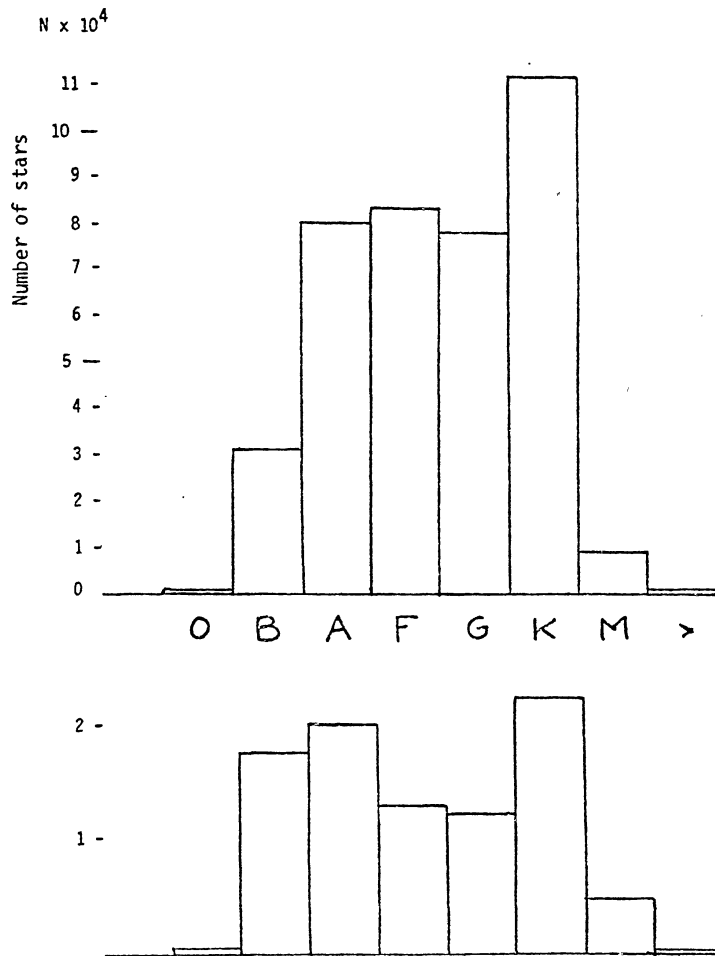


Figure 2 : Spectral type distribution of the CSI stars (top) and of the stars brighter than V=6.5 (from the Catalog of Bright Stars, Hoffleit, 1982) (below)

...

3.5 The cross-identifications: the CSI is primarily a dictionary of identifications: the identifications (or running numbers) from the catalogs listed in Table 1 are included. Two examples (Vega and Arcturus) are presented in Figure 3. Note that not only usual names and identifications -such as Vega, α Lyr, HR 7001, ADS 11510 A- are included, but also running numbers in compilations of observational data, such as UBV 15842 or UVBY 4430, referring to the 15842nd or 4430th star in the corresponding file (there, the star is listed under one of its usual identifications, namely the HD number or its equivalent in the Mermilliod numbering system).

```

* ----HD 124897 -----CSI +19 2777 C-----
COORD. 1950. = 14 13 22.750+19 26 30.96          SP = K2 IIIP
              PM = -1.098 [ 1 ] / -1.999 [ 1 ]      MB,P,MV = 1.19 -0.04

BD +19 2777      HD 124897      * 16 RCO      * ALF BCO
NAME ARCTURUS   FE 251          FK4 526       GC 19242
GCRV 8341       GEN 2096       GJ 541        HR 5340
IRC +20.270     JP11 2486       MKJ 8632      N30 3229
PLX 3242        ROT 2044       SAO 100944    UVBY 12551
UVBY 3398

-----
*MEASUREMENTS :
MKJ (32)  UBV ( 7)  GEN ( 1)  JP11( 4)  IRC ( 1)  UVBY( 3)
PLX ( 5)  GCRV(41)  ROT ( 2)  SAO ( 1)  N30 ( 1)  GC ( 1)
GJ ( 1)

-----
357 BIBLIOGRAPHICAL REFERENCES
-----

* ----HD 172167 -----CSI +38 3238 1-----
COORD. 1950. = 18 35 14.655+38 44 9.6E          SP = AC VVAR
              PM = +0.260 [ 1 ] / +0.285 [ 1 ]      MB,P,MV = 0.03 0.03

BD +38 3238A    HD 172167      * 3 Lyr      * ALF Lyr
NAME VEGA      ADS 11510A     AG +38 1711  CEL 4636
FE 345         FK4 699       GC 25466    GCRV 11065
GEN 2441       GJ 721        HGAM 706    HR 7001
IRC +40.322    JP11 2999     MKJ 10485   N30 4138
PLX 4293       ROT 2633     SAO 67174   LBV 15842
UVBY 4430      1E 183515 +3844.3  CSI +38 3238 1

-----
*MEASUREMENTS :
MKJ (33)  UBV (13)  GEN ( 1)  JP11( 5)  IRC ( 1)  CEL ( 1)
UVBY( 6)  FLX ( 4)  GCRV(16)  ROT ( 2)  HGAM( 1)  SAO ( 1)
AG ( 1)  N30 ( 1)  GC ( 1)  GJ ( 1)

-----
465 BIBLIOGRAPHICAL REFERENCES
-----

```

Figure 3 : Examples of fundamental data and lists of identifications available for Arcturus and Vega. The numbers of measurements and bibliographical references also available in the data base are given (SIMBAD - March 1983)

In the original organization of the CSI, only identification flags were stored in the main file, the cross-identifications (DM number → flagged identification) being stored in a second file (ID). This file is now merged with the main SIMBAD file, on the Strasbourg site. Access to the Data Base is gained through the reverse file ID⁻¹ by using *any of the existing names or running numbers*. It is clear that this facility, offered by the CSI, is a very important improvement in the field of astronomical data. This facility is being extended to all the designations appearing in the literature and collected for the BSI (4.2). About 220 catalogs or lists, used in the literature, are presently taken into account (1.9×10^6 identifications).

3.6 The multiple systems: an important point is the treatment of multiple systems. For the purpose of precise cross-identification, it appeared useful to distinguish whether or not the components had been observed separately in a given catalog. As a consequence A and B components may be included both as separate entries (for instance, from an astrometric catalog) and the AB system as an additional entry (related, for instance, to the measurement of the system as a whole in a photometric catalog). In such a case, a double system may have three entries in the CSI. In the whole CSI about 1000 additional entries (indicated by a flag) refer to combinations of stars appearing separately elsewhere in the catalog.

Example :	CSI key number	Other identification
	+9 1064 1	ADS 4617 AB : BD +9° 1064, HR 2124, etc.
	+9 1064 2	ADS 4617 C
	+9 1064 3	ADS 4617 A
	+9 1064 4	ADS 4617 B

Note : the code (1, 2, 3, 4) in the CSI key number discriminates the separate entries independently of the classification (A, B, C, etc.) in the catalogs of multiple stars.

4. THE OBSERVATIONAL AND BIBLIOGRAPHICAL DATA

In the previous organization of the CSI on the Meudon site, the identification (or running number) of the star in a number of observational catalogs gave access to the corresponding data stored on direct-access files (file MES); the bibliographical data from the BSI were also available through the CSI for the stars common to the two (for more details see Ochsenbein et al., 1977 and Ochsenbein, 1981). Now, on the Strasbourg site *both kinds of data are included* in the main SIMBAD file.

4.1 The observational measurements: the immediate access to the observational data does not exist for all the connected catalogs (and, a fortiori, for all the catalogs existing on tape at the Stellar Data Center), but only for the major compilations and fundamental catalogs (see Table 1). This access generally concerns only the specific data in the given catalog: for instance immediate access is given to the MK classifications of the MK Classification Extension (Kennedy, 1978) and their bibliographical references, but not to the magnitudes and coordinates provided by Mrs Kennedy on the magnetic tape (which are, generally not original data). For the AGK2/3 and the GC, only proper motions (and not the coordinates) are stored while both kinds of data are stored for the N30 and the SA0.

A catalog is generally not 100% connected to the CSI: to begin with, the only stars included are those for which a cross-identification exists with a list previously connected to the CSI. Later, new cross-identifications are searched, case by case, in the literature, and multiple systems are also examined in detail. And still later, new stars (not appearing in a previously linked catalog) are added to the CSI: this last step is possible only when coordinates are known with sufficient precision, and is not achieved (nor

achievable !) for all the catalogs. An index of the cross-identifications found for the 22 700 non-DM faint stars included in the CSI will be published in the near future (Bischoff, 1983). The percentage of stars at present included in the CSI for the connected catalogs is given in Table 1. Connections of other catalogs are projected.

The surface distributions of some basic observational data were given some years ago by Egret et al. (1976). But there is constant progress in numerous fields: photometry, MK classification, radial velocity, etc. With the help of collaborating institutes the CDS is keeping the major files of observational data up-to-date (see Jaschek, 1977; and regular reports in the CDS Information Bulletin). New suggestions for catalogs that it would be useful to include in the Data Base are welcome.

With regard to the observational measurements, the Data Base always provides the bibliographical reference of the original publication or of the compilation catalog from which the data are issued. Homogenized or averaged values are generally given when available, but also the original measurements: the user is always able to evaluate the quality and the dispersion of the data. He may also choose to use the mean (or "best") value established by the specialists in the field (but not by the CDS on its own).

4.2 The bibliographical data : the BSI gives access to the bibliography of the star from 1950 on; the complete reference and title of any paper in which the star is cited is recorded (provided that this paper appeared in one of the major journals surveyed). It has been described by Ochsenbein and Spite (1977) and Ochsenbein (1981, 1982). The journals surveyed are listed at the beginning of the published microfiches covering the periods 1950-1972, 1973-1977, and 1978.

The BSI is constructed as a separate catalog and includes all the stars cited in the papers covered by the survey (lists of more than 1000 objects are excluded and considered as catalogs). But the connection with the CSI gives the enormous advantage, for the stars common to the two, that one can consider all the citations of a star, whatever identification is used: HR, HD, ADS, DM, etc. (see Table 1). The BSI's reliability has been discussed by Nishimura (1982). The last published edition of the BSI, for the period 1950-1978 includes 350 000 citations (referring to 89 000 different stars) from 12 000 papers (Ochsenbein, 1982).

5. ORGANIZATION OF THE DATA BASE

Originally, on the Meudon computer, the Data Base comprised the CSI and the BSI plus the files ID⁻¹, ID, and MES (see e.g., Fig. 2 of Ochsenbein et al., 1977). The organization of SIMBAD, on the present Strasbourg site, is very different with regard to data processing: the files CSI, ID, MES and BSI are now merged in a single one, with one record per star, but of variable length. This record contains the data described in the previous sections, namely: fundamental data, star identifications, observational measurements (and related references), and bibliography. In SIMBAD (500 000 entries at the end of 1982) some of the stars do not figure in the CSI, but only in the BSI or an incompletely connected observational catalog (previously file MES). Such stars do not have "CSI numbers" and their coordinates are not known; this is, in fact, a temporary arrangement, and they will be integrated gradually into the CSI as soon as coordinates and/or cross-identifications are found (see 4.1 and the catalog in preparation by Bischoff M., 1983).

Use of the Data Base is possible either in batch mode, or in interactive mode using a video display terminal (which may be associated with a printer). The access to the Data Base is possible through any of the recorded names, identifications (from the

CSI or the BSI), or running numbers in a connected catalog, existing for a star. The interactive procedure allows subsets of the data to be selected. Generic interrogations are also possible (e.g.: SAO 1 to 100), and sampling procedures, by astrophysical parameters and coordinates, are being developed (Wenger, 1981, and further announcements in the CDS Information Bulletin). For more practical details the reader should refer to the "SIMBAD user's guide" prepared at Strasbourg by M. Wenger. Other institutes can link directly with SIMBAD using the existing networks (Wenger, 1980). Remote-access terminals exist presently in Madrid-Villafranca (ESA Station), Paris-Meudon, Paris (IAP), Nice, Marseille (LAS), Grasse (CERGA), and Besançon. Other connections are projected.

6. FILES AND CATALOGS DERIVED FROM THE DATA BASE

SIMBAD is a complex set of files, stored in UNIVAC-dependent binary format. It is not possible to transfer the whole as such to other computers. It is therefore necessary to extract files which are subsets of the Data Base, edited in a convenient format, and, what is more, adapted to answer a given problem. I give here some examples of such files.

6.1 The 1979 edition of the CSI: the most recent microfiche edition of the CSI, described by Ochsenbein et al. (1981) consists of three parts: the CSI itself (fundamental data and identifications) on 46 microfiches, sorted by declination zones and DM numbers (the non-DM stars being incorporated according to their right ascension); the cross-reference tables (ID^{-1}) providing an entry to the CSI through one of the following designations: HD, HR, ADS or IDS, name, or variable name (7 microfiches); selected data from the main catalogs sorted by catalog number (26 microfiches). This edition is convenient for examining a limited number of objects. The first part is also available on magnetic tape (CDS reference 4009): it is a copy of the microfiche edition, not very convenient for computer processing because it was not possible to allocate one fixed-length column for each possible identifier. Copies of the second and third parts may also be made for special requests.

6.2 The SAO Catalog with Astrophysical Data : this subset of the Data Base gives selected data for all the SAO stars. Quite a number of errors in the original edition have been corrected (Bischoff, 1978, for instance), and the catalog includes 258 968 stars (258 997 in the original catalog). This edition, published on microfiche (Ochsenbein, 1980) contains the following data: SAO number, HD number, DM (or CSI) number, coordinates for equinox and epoch 1950.0 (from the SAO), galactic coordinates, proper motions and quality index, UBV photometry when available or else CSI magnitudes (see 3.3), MK spectral type when available or else HD type (3.4), radial velocity, and remarks (detected errors, trigonometric parallax, double star designation, and name) (26 microfiches). Two additional microfiches provide an HD/SAO index. UBV photometry, MK type, and radial velocity included are selected values from the catalogs by Nicolet (1978), Jaschek (1978), Houk and Cowley (1975), Houk (1978), Evans (1967), and Wilson (1953).

A copy of the microfiche edition is also available on magnetic tape (CDS reference 5015); but, here again, this tape is not convenient for computer processing (headings and remarks are inserted among the data records).

6.3 CSI Catalog with Selected Data (1982): this catalog, recently made available on magnetic tape (ref. 5026), includes positions, proper motions, UBV photometry, radial velocity, and MK classification for 434 023 stars. The only identification given

is the CSI key number (3.1 above), and the file is sorted by coordinates. UBV, MK, and RV data have the same source as in the SAO Catalog just described (6.2). Proper motions come from the SAO and the AGK3. Additional parameters were estimated, for the purpose of the preparation of a list of stars expected to be nearer than 200 pc (Ochsenbein and Egret, 1981): absolute magnitude (from Gliese's Catalog or estimated spectroscopically) and visual interstellar extinction (computed from spectroscopic and photometric data). A coded version of the MK classification, used in the computations, is also recorded. All these data (including bytes providing the source) are recorded in a single fixed-length (144 characters) record per star. A paper describing the method and calibration tables used for estimating M_V and A_V is in preparation (Ochsenbein F., 1983) When no estimation was possible the field is left blank. All the stars in the 1979 edition of CSI are included, except the additional entries in multiple systems (see 3.6).

I wish to add here a personal remark: experience with some catalogs such as the SKYMAP (Gottlieb, 1978) has shown that great care should be taken when publishing *estimated* data: these data are, by definition, debatable and they become useless for the reader as soon as he does not fully agree with the calibration tables used, not only useless but also dangerous if some fields become filled with "extremely unreliable" derived parameters.

6.4 The preparation of the HIPPARCOS Input Catalogue : an example of a possible file derived from the Data Base may be taken from the preparation of the Input Catalogue for the ESA astrometric satellite HIPPARCOS. The Stellar Data Center is a member of the Input Catalogue Consortium (INCA) and will participate by providing a file which may be used as a framework for the Input Catalogue, to be filled out by additional data that remain to be collected (Egret, 1982).

The stars to be included come from: 1) submitted proposals; if the identification used is included in the CSI, it is easy to find the corresponding data in the Base (see above, section 5); 2) a general survey of all stars brighter than 8.5 magnitudes (or some other limit to be defined): the Data Base is then used as a sequential catalog and is certainly complete to this extent.

For each selected star the following data will be extracted: at least one identification, coordinates (from the CSI or from an astrometric catalog: see 3.2), magnitudes (3.3), spectral classification (3.4), and proper motions (3.2). Secondly, complementary data about the multiplicity and variability of the star (using the ADS, the IDS, the GCVS) will be taken from the Data Base. Thirdly, additional astrophysical data can be considered, such as trigonometric parallaxes (if any), radial velocities, etc. The data mentioned above can easily be stored on a magnetic tape containing one fixed-length record of approximately 200 bytes per star.

For a number of objects (especially the fainter ones), some of the data are not existing in the Base; some of the stars proposed might also not appear in the Base: the INCA group will undertake the necessary investigations and, if necessary, make new measurements or observations. The CDS intends to include all these future data gradually, as well as possible corrections related to the present data.

7. THE COMPLETENESS OF THE CSI

Let us, finally, examine the completeness of the CSI: to what limit of magnitude can the CSI (or the Data Base) be said to be complete? The answer is not simple. The completeness limit results from the combined limits of the basic catalogs used for creating the Data Base. But this limit differs greatly from one cata-

log to another: most surveys do not cover the whole sky; some are photographic and others are visual, and the corresponding magnitudes may differ greatly from the photoelectric UBV magnitudes (Bischoff and Ochsenbein, 1983). Other catalogs (the SAO) apply a density limit which excludes a number of rather bright stars. On the whole we can estimate that the CSI includes *all* stars brighter than $V = 9.0$ (and probably the limit of completeness is even slightly fainter). But for a number of reasons (precision of photometric data, variability of stars, clerical errors of all kinds) the word "all" in the preceding sentence means essentially "more than 98 or 99%". A study of the completeness of the data included in the Data Base is planned (Egret, 1983)

8. REFERENCES TO PAPERS CONCERNED WITH THE CSI AND SIMBAD

Reference	Subject
Jung, J. (1971)	Report on the Strasbourg Stellar Data Center
Jung, J. and Bischoff, M. (1971)	The General Catalogue of Stellar Identifications (CSI)
Jung, J. and Ochsenbein, F. (1972)	The General Catalogue of Stellar Identifications: A) Progress Report - B) Statistics Pertaining to the Overlap between Catalogues Included in the CSI
Jung, J., Bischoff, M. and Ochsenbein, F. (1973)	The Catalogue of Stellar Identifications (CSI): A Progress Report II
Jung, J. (1973)	A Statistical Description of the Catalogue of Stellar Identifications (CSI) - 1
Ochsenbein, F. (1974a)	On the Relationship between the Apparent Magnitudes Given in Several Catalogues and the UBV System
Ochsenbein, F. (1974b)	Automated Information Retrieval at the Stellar Data Center
Ochsenbein, F., Bischoff, M. (1975)	The Catalogue of Stellar Identifications: Progress Report III
Ochsenbein, F. (1976)	The Present Status of Data Available at the Stellar Data Center
Egret, D., Heck, A., Ochsenbein, F. (1976)	Surface Distributions of Some Basic Observational Data
Ochsenbein, F., Egret, D., Bischoff, M. (1977)	The Catalogue of Stellar Identifications
Jaschek, C. (1977)	The Strasbourg Stellar Data Centre
Ochsenbein, F. (1978)	The Catalogue of Stellar Identifications: Main Features
Ochsenbein, F. (1979)	Automated Information Retrieval System: Presentation of the Data
Egret, D. (1979)	Uses of Catalogues of Data in the Stellar Data Center
Ochsenbein, F., Bischoff, M., Egret, D. (1981)	Microfiche Edition of CSI
Wenger, M. (1981)	Base de Données du CDS : SIMBAD
Ochsenbein, F. (1981)	Le CSI
Nagy, T.A., Hill, R.S., Mead, J.M. (1981)	CSI Statistics: Blue Magnitude versus Spectral Type
Egret, D. (1981)	The Strasbourg Stellar Data Center
Ochsenbein, F., Bischoff, M. (1982)	The Catalogue of Stellar Identifications
Ochsenbein, F. (1983)	CSI Catalogue with Selected Data

9. CONCLUSION

The first aim of this paper is to encourage our colleagues to make extensive use of the CDS Data Base. The Staff of the CDS will gladly help. A number of conversational procedures are presently being developed. They make the Base easily accessible to persons not familiar with programming.

New developments are being prepared and will be announced in the CDS Information Bulletin: inclusion of recent photometric and spectroscopic data, addition of data on variable stars, etc.

Another extension, towards galaxies, is under way (Dubois et al., 1983, in the same issue of the Bulletin).

Acknowledgements. I thank the colleagues at the CDS who answered a number of troublesome questions during my preparation of this paper: Mme M. Bischoff, F. Ochsenbein, and M. Wenger. The development of the Data Base has been a collective undertaking to which also J. Jung (former Director), C. Jaschek (Director of the CDS), and Mme M.-J. Wagner (who did most of the keypunching) are associated.

REFERENCES

- Bischoff, M. : 1978, Bull. Inform. CDS 14, 2
- Bischoff, M. : 1983, in preparation
- Bischoff, M. : 1983, Bull. Inform. CDS 24, 47
- Dubois, P., Ochsenbein, F. and Paturel, G. : 1983, Bull. Inform. CDS 24, 125
- Egret, D. : 1979, Dudley Obs. Reports 14, 393
- Egret, D. : 1981, in "Answer to the Announcement of Opportunity for the Compilation of the HIPPARCOS Input Catalogue", C. Turon ed., p. 239
- Egret, D. : 1982, Bull. Inform. CDS 22, 63
- Egret, D. : 1983, in preparation
- Egret, D., Heck, A. and Ochsenbein, F. : 1976, Astron. Astrophys. Suppl. Ser. 26, 65
- Evans, D.S. : 1967, in IAU Symp. 30 "Determination of Radial Velocities and their Applications", A.H. Batten and J.F. Heard, eds., p. 57
- Gottlieb, D.M. : 1978, Astrophys. J. Suppl. Ser. 38, 287
- Hoffleit, D., with the collaboration of Jaschek, C. : 1982, "The Bright Star Catalogue", 4th revised edition, Yale University Observatory
- Houk, N. and Cowley, A.P. : 1975, "Michigan Catalogue of Two-Dimensional Spectral Types for the HD Stars", Vol. 1, Univ. of Michigan
- Houk, N. : 1978, "Michigan Catalogue of Two-Dimensional Spectral Types for the HD Stars", Vol. 2, Univ. of Michigan
- Jaschek, C. : 1977, Vistas in Astronomy 21, 311
- Jaschek, M. : 1978, Bull. Inform. CDS 15, 121
- Jaschek, C., Ochsenbein, F., Bidelman, W.P. : 1980, Bull. Inform. CDS 18, 41
- Jung, J. : 1971, Bull. Inform. CDS 1, 2
- Jung, J. : 1973, Bull. Inform. CDS 5, 1
- Jung, J. and Bischoff, M. : 1971, Bull. Inform. CDS 2, 8
- Jung, J. and Ochsenbein, F. : 1972, Bull. Inform. CDS 3, 6
- Jung, J., Bischoff, M., Ochsenbein, F. : 1973, Bull. Inform. CDS 4, 27
- Kennedy, P.M. : 1978, "MK Classification Extension", Mt Stromlo Obs.
- Lacroute, P. and Valbousquet, A. : 1971, in IAU Coll. 7 on "Proper Motions", W.J. Luyten ed.
- Nagy, Th. A., Hill, R.S., Mead, J.M. : 1981, ADC Bull. 1, 121
- Nicolet, B. : 1978, Astron. Astrophys. Suppl. Ser. 34, 1
- Nishimura, S. : 1982, Bull. Inform. CDS 22, 90
- Ochsenbein, F. : 1974a, Astron. Astrophys. Suppl. Ser. 15, 215
- Ochsenbein, F. : 1974b, Bull. Inform. CDS 7, 7
- Ochsenbein, F. : 1976, in "Proceedings of the Third European Astronomical Meeting", Tbilissi, p. 497
- Ochsenbein, F. : 1978, Bull. Inform. CDS 15, 88
- Ochsenbein, F. : 1979, Bull. Inform. CDS 16, 105
- Ochsenbein, F. : 1980, Bull. Inform. CDS 19, 74
- Ochsenbein, F. : 1981, Thesis, Université Louis Pasteur, Strasbourg, pp. 9-18
- Ochsenbein, F. : 1982, in IAU Coll. 64, "Automated Data Retrieval in Astronomy", C. Jaschek and W. Heintz eds., p. 171
- Ochsenbein, F. : 1983, in preparation
- Ochsenbein, F. and Bischoff, M. : 1975, Bull. Inform. CDS 8, 2
- Ochsenbein, F. and Bischoff, M. : 1982, in IAU Coll. 64, "Automated Data Retrieval in Astronomy", C. Jaschek and W. Heintz eds., p. 211
- Ochsenbein, F. and Egret, D. : 1981, Bull. Inform. CDS 20, 60
- Ochsenbein, F. and Spite, F. : 1977, in IAU Coll. 35, "Compilation, Critical Evaluation and Distribution of Stellar Data", C. Jaschek and G.A. Wilkins eds., p. 175

- Ochsenbein, F., Egret, D. and Bischoff, M. : 1977, in IAU Coll. 35, "Compilation, Critical Evaluation and Distribution of Stellar Data", C. Jaschek and G.A. Wilkins eds., p. 31
- Ochsenbein, F., Bischoff, M. and Egret, D. : 1981, Astron. Astrophys. Suppl. Ser. 43, 259
- Wenger, M. : 1980, Bull. Inform. CDS 19, 20
- Wenger, M. : 1981, Bull. Inform. CDS 20, 81
- Williams, P.M. : 1967, Mon. Notes Astron Soc. South Africa 26, 30
- Wilson, R.F. : 1953, "General Catalogue of Stellar Radial Velocities", Carnegie Institution of Washington, Publ. 601

Author's address : D. EGRET
Observatoire de Strasbourg
Centre de Données Stellaires
11, Rue de l'Université
67000 STRASBOURG - France