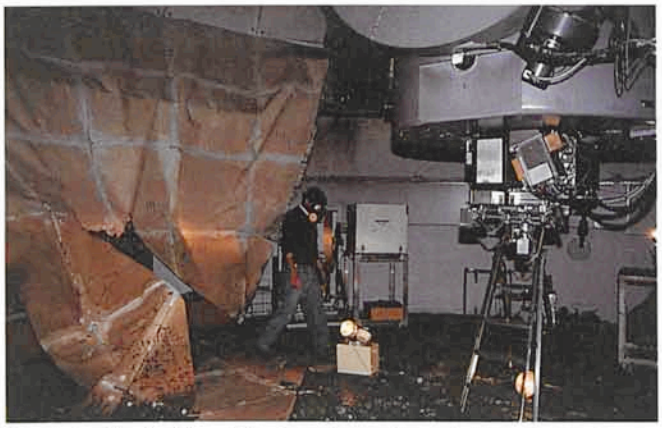

Figure 2: *The inside of the 1-m dome, just after the fire had been extinguished. The 1-m telescope and the multichannel photometer fortunately do not seem to have suffered bad damage. This was later confirmed.*

Figure 3: *On the next day, the cleaning work began. Thanks to a very dedicated effort by ESO staff, this only lasted two weeks and the telescope was again in operation by mid-November.*

scope control system suffered no major damage. The mirrors had to be cleaned; the secondary was realuminized, while it is not necessary to realuminize the primary immediately.

The image quality fortunately did not suffer from this bad experience, as shown by an Antares run conducted on November 6. The instruments on the telescope – a special photometer just installed when the fire started – did not suffer much, although it obviously had to be cleaned very thoroughly.

The La Silla operations and mechanics and construction groups made a tremendous effort to quickly return the telescope to the astronomical community. The normal schedule of observations started again on November 12, following some necessary test nights. *A. SMETTE, ESO-La Silla*

Astronomical Data Handling: Windows of Opportunity and of Challenge

F. MURTAGH, ST-ECF, Garching

It is the destiny of astronomy to become the first all-digital science. [1]

1. Medium and Message

The different sections of this article start up a few windows, in order to view some current developments affecting astronomical storage and retrieval. As an area driven powerfully by collective research efforts, we can only offer a small (but colourful!) palette of what is currently available.

Dusty card decks, screaming paper tape and (soon) cumbersome 9-track magnetic tape reels have given way to storage devices of the sort shown in Figure 1. The reel of 9-track $\frac{1}{2}$-in tape is shown for comparison.

Magnetic media include QIC, Exabyte and DAT cartridges, respectively using 6 mm, 8 mm and 4 mm tape technology.

QIC (referring to "Quarter Inch Cartridge", rather than speed) is similar to 9-track tape in capacity, but several times more compact. This streamer tape storage medium potentially suffers from heat, static build-up, and resultant positioning difficulties. Exabyte and DAT are helical-scan (hence storage-efficient) tape devices. The former is marketed uniquely by Exabyte. Unit costs for these storage media are uniformly low.

Optical media are laser-read and hence less susceptible to mishaps like head-crashes. They also are unhindered in backspacing. Periodic head realignment may however be necessary. Here we will comment on compact disks, optical disks both read-only and read-writable, and optical tapes.

CD-ROMs ("Compact Disks – Read-Only-Memory") are of somewhat better quality than their audio (music) siblings. They are being increasingly used for storage of astronomical catalogues (somewhat disingenuously referred to as "dead data"). The 12-in diameter optical disk shown in Figure 1 is of the sort used for receipt of Hubble Space Telescope archive data in Europe. The storage supported by CD-ROMs is soon to be 4 GB. And what of the near future? We will see widespread usage of erasable, or read-write, optical disks, using magneto-optic (MO) technology. These will, inter alia, come in $3\frac{1}{2}$-in diameter sizes, will cater for 128 MB, and will be attachable to anyone's laptop or notebook. For larger-scale applications,

[1]Larry Smarr, University of Illinois, quoted in *The Economist*, October 17, 1992.

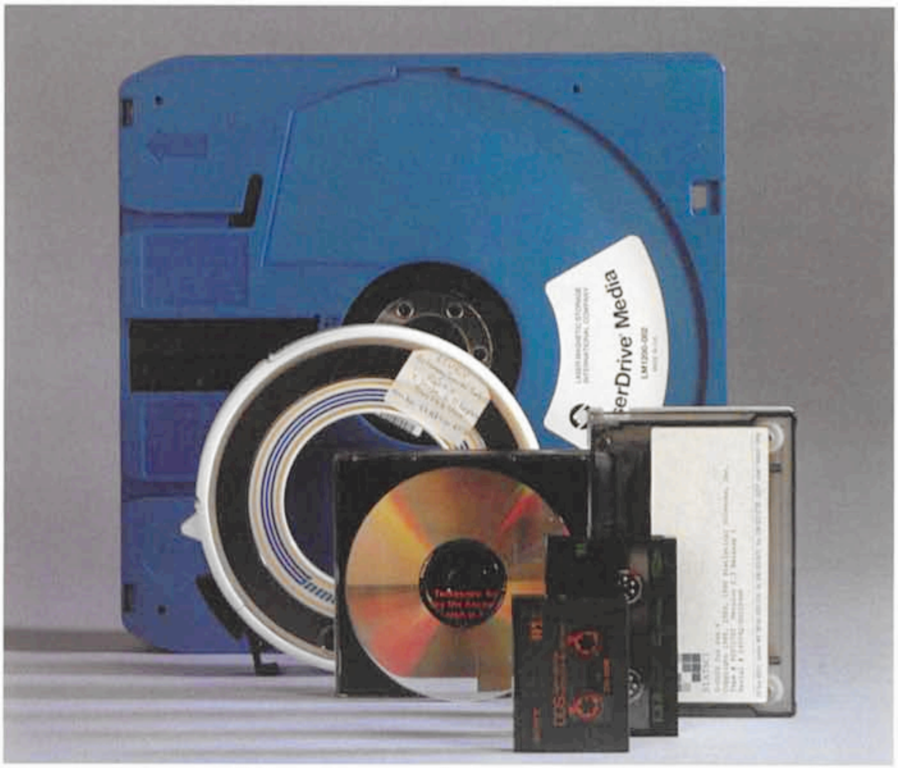

Figure 1: *Optical disk; 9-track tape (for comparison); CD-ROM; QIC; Exabyte; DAT (front).*

optical tape will allow for storage of 1 TB. Juke-boxes, to expedite the mounting and unmounting of storage devices, are still very expensive.

The examples of storage media discussed here refer to a very small range of what is available in the market-place. When it comes to the transfer of data, one should keep computer network in the pictures also. Finally, as the storage standard in common use, FITS ("Flexible Image Transport System") has imposed itself everywhere – an indication of unequivocal success.

2. Beyond WYSIWYG: What You See Is What You Want

A powerful impulse to changing the way we interface with machines was the introduction of the Macintosh by Apple in 1984. Its windows, icons and pull-down menus launched a thousand graphical user interfaces.

Support for multitasking and support of graphics are among the machine qualities that we now take for granted. The X Window System arose out of MIT's Project Athena, a joint project between MIT, DEC and IBM which began in May 1983. The current version of X, X11, has become the most widely-used windowing system on scientific workstations. *Motif* is one among many interface systems which have been built on top of X.

More basic prompt and menu-driven user interfaces are still widely used. For example, one can access the NSSDC On-Line Data & Information Service at Goddard by telneting to *nssdca. gsfc. nasa.gov* and using *nodis* as the username. But the user will generally have to wait until ordered information shows up, before accessing its relevance.

Table 1: *Indicative storage capacities and costs of drives for a range of media.*

Device	Indicative max. capacity	Indicative cost of drive
QIC	250 MB	DM 1000
DAT	2 (8) GB	DM 3000
Exabyte	2.5 (5) GB	DM 5000
CD-ROM	650 MB	DM 1000
optical disk	1.2 or 3.2 GB	DM 50000
optical tape	1 TB	DM 200000

Notes on Table: Indicative maximum capacity figures in parentheses are with use of compression; optical disk figures are for one-sided use, for different vendors. 1 MB is 1×10^6 bytes (a character or a number), 1 GB is 1×10^9 bytes, and 1 TB is 1×10^{12} bytes).

Quick-look capability introduces the same sort of inestimable improvement into working with databases as interactive computing brought to the world of batch computing. Images can be quite large (e.g. around 10 megabytes for some HST images), thereby slowing down any quick-look implementation. The clever scheme implemented in STARCAT is based on a compressed version of the image. Wavelet-based compression is one approach which has proved very effective. If loss of information is acceptable as is the case for quick-look information, then $\frac{1}{20}$th of the original image's size may suffice for an acceptable visual interpretation. The slimmed-down, but faithful, representation of the image lets the user know in real time if the image is worth retrieving or not. Figure 2 shows an example of the use of the STARCAT preview facility. Whether used for spectra or 2-d images, the potential savings in time and effort on the part of the user can be enormous.

3. United Colours of Astronomy

Panchromatic or pluri-disciplinary astronomy implies the use of various wavelength windows and of the results of various observational instruments. It comes at a price. Diverse data collections can occasion ambiguity, confusion and error. Among pitfalls are: physical location of data, and access procedures; nomenclature and content-characterization; homogeneity and reliability; and so on. The user must first identify what catalogues or databases are relevant; secondly, appropriate access conventions must be availed of; and thirdly, special analysis utilities may be required in order to combine different data.

The latter aspect may involve graphical approaches – image overlays and mosaics, for example; or regression or other data summarizing methods using data values affected by various errors or censoring; or model-driven data fusion such as is practised in multiframe image restoration.

Support for some of these options is becoming possible with the ESIS ("European Space Information System") Correlation Environment (CE). Figure 3 shows a summary statement of what is available on an object, across various databases. A bibliographic survey (using SIMBAD) is also illustrated.

Beyond this, the CE is ambitiously designed to be an evolving collection of display and manipulation tools. The CE is disciplined by connectivity to a query environment which is more precise and limited in its functionality.

With the CE, the "correlation of data", we are entering new and uncharted

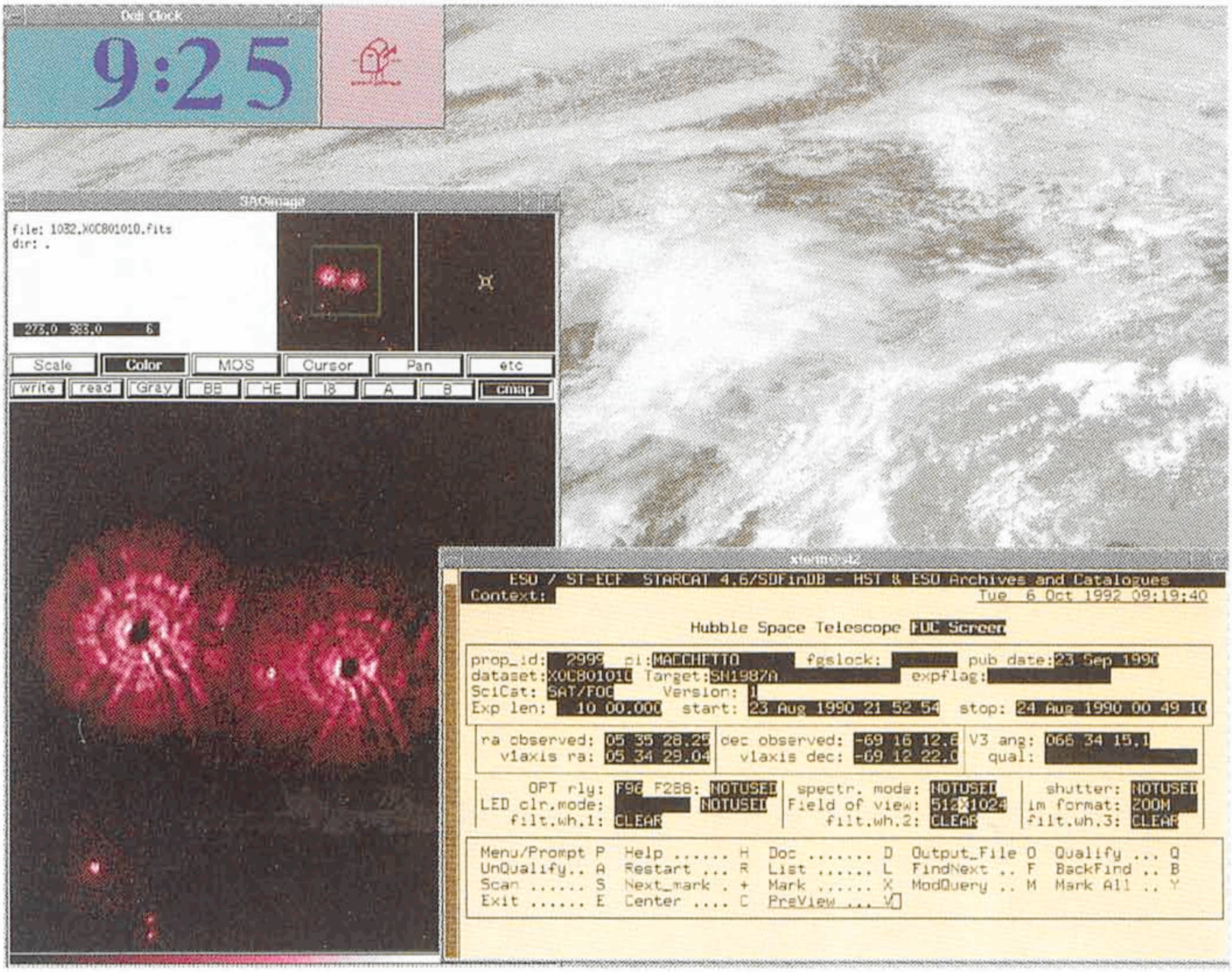

Figure 2: *Previewing an HST Faint Object Camera image of 1987A using STARCAT.*

territory. It is, though, the type of system which is needed to channel the ever-growing flood of observational data, and to satisfy the widely-understood need that expensively-gathered data is a valuable resource for all who wish to avail of it.

4. Literature

Literature has, in time-honoured fashion, been considered a humanizing influence. How does astronomy fare in this regard? More on the point, *where* is the literature when you need it within seconds, if not sooner?

The widely-used SIMBAD database contains bibliographic references – up to 600 for an individual object. It contains approximately 1 million references in total, covering about 60,000 papers, derived from 90 astronomical periodicals.

Abstracts may be accessed through the *Astronomy and Astrophysical Abstracts* (AAA), formerly from 1899 to 1969 under the title of *Astronomischer Jahresbericht*. AAA is available since 1985 in computer-readable form. It is now available in the PHYS on-line bibliographic database supported by the FIZ (Fachinformationszentrum für Physik, Karlsruhe).

Titles and abstracts help, but evidently fall short of what could be possible with full-text retrieval. The copyright situation casts a long shadow over what is technically feasible and – on the part of the end-user – desirable. Preprints, though, are usually in a sort of benevolent zone. Their transitory nature allows them to partially escape from otherwise stringent proprietary controls. The value of having preprints on-line is inestimable for three reasons: (i) savings of paper and postage, (ii) speeding-up of the transfer of information to the scientific consumer, and (iii) opening up of the possibility of using full-text retrieval mechanisms.

A line-mode preprints server, covering astronomy as well as other areas of physics, is run by the International School for Advanced Studies in Trieste. Access by mail is supported, as is a mail-based alert system for new contributions, and anonymous ftp (to *babbage.sissa.it*). Thus far, the number of astronomy-related papers is not large.

The preprint situation at CERN is inspirational. The CERN printshop outputs 60 million A4 pages per year, of which 20% are preprints. Standardization on printer systems which adhere to the PostScript standard has reduced production time of preprints from weeks to days. Spurred on by speedier appearance, submission of preprints in PostScript by authors rose from 5% to 60% in one year. PostScript is, of course, the page description standard supported by most text processing systems (TeX and LaTeX, MS Word, PageMaker, WordPerfect, most graphics packages, etc.). The next step was to further reduce delays from scientific producer to scientific consumer by simply having permanent on-line access to all preprints. Figure 4 shows an example of the use of the X-Windows interface to this server. Preprints can be pulled over at the proverbial drop of a hat. A simpler line-mode interface is also supported.

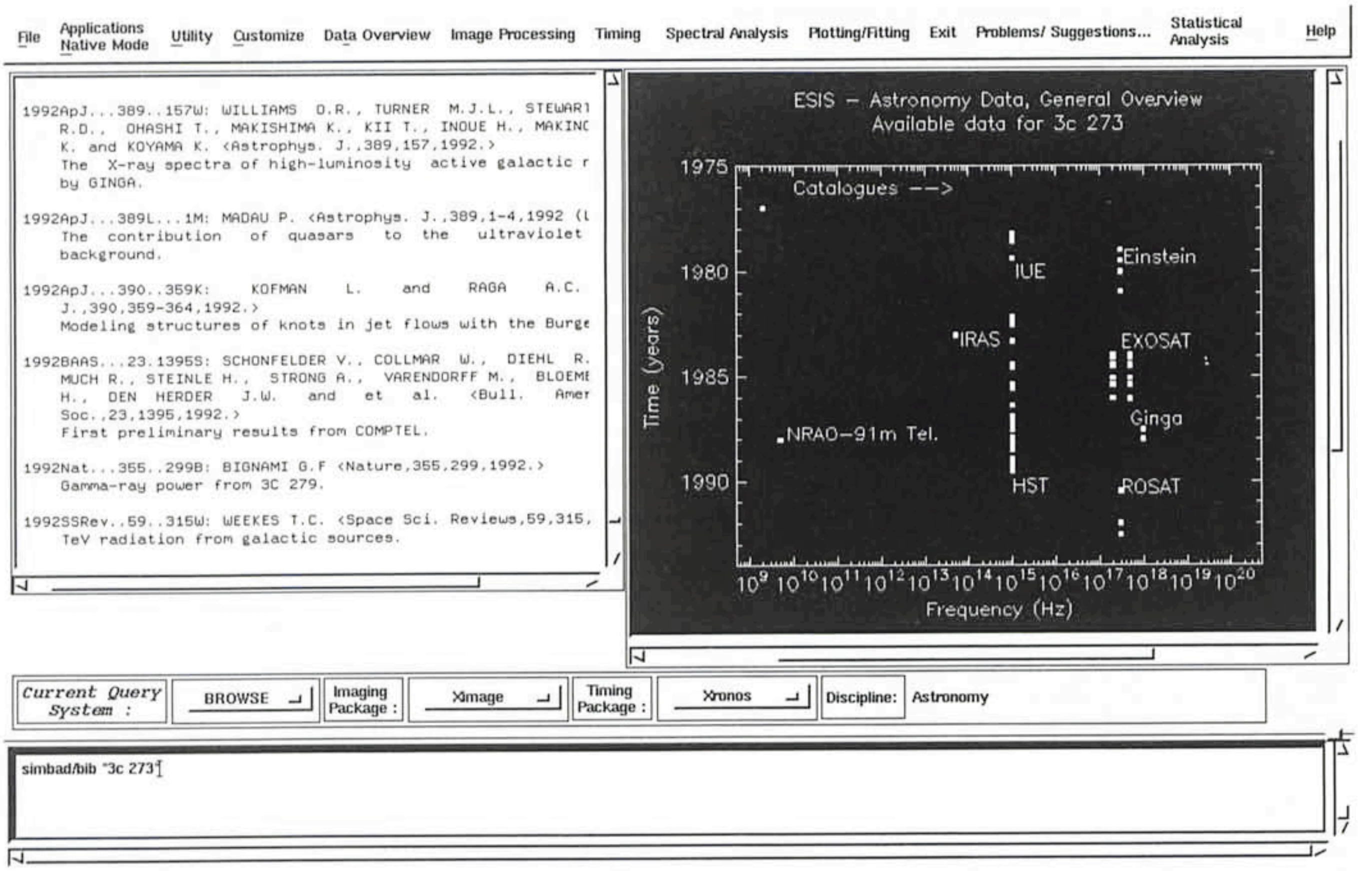

Figure 3: *Matching data on 3C 273 from various databases, and from the literature, using ESIS's Correlation Environment.*

5. The Late 20th Century's Lingua Franca

We may have different national origins, but we all now speak one language: tpc/ip. A mild exaggeration perhaps, not least because various other computer communication protocols can make themselves understood loudly and clearly to the tcp/ip protocols. Furthermore middle-level, hand-shaking, data-transferring protocols such as tcp ("transmission control protocol") and ip ("internet protocol") are not as user-relevant as even higher-level protocols. In this section, we will look at what is possible using some of these higher-level standards.

The basis of fact in this section's title is that the Internet is closely related to these protocols. The internet is made up of a large number of backbone, regional and local data networks.

The Usenet/Internet newsgroup system presents a daunting and fascinating communication subsystem. An estimated 60,000 sites have access to the 8000 newsgroups and about 1.6 million users at these sites are newsreaders. The approximate 16,000 message postings per day correspond to over 40 megabytes of traffic. Of course the signal-to-noise ratio is not always high: today's undergraduate students get their term papers done by posting a question, and waiting for the answers to roll in.

Cleverly, though, it has been noted that even low-grade questions can play a beneficial role. In many newsgroups, a FAQ (Frequently Asked Questions) posting is regularly updated. The FAQ (including answers) for the astronautic newsgroup, sci. space, for instance, comprises 170 kilobytes. A FAQ is a paradigmatic dynamic document, which collects many of the most frequent questions and answers. FAQs are pedagogical documents par excellence and can be availed of by anyone seeking a quick grounding in a particular subject.

The common data transfer protocol (and command name) on the Internet is ftp ("file transfer protocol"). Anonymous ftp has become a standard way to make datasets, documentation, and code available.

A qualitative step forward is currently taking place in regard to what is available through anonymous ftp. Indexing and stock-taking of what is available is increasingly being carried out in an automated way. The archive server, archie, regularly polls 1000 anonymous ftp sites, seeking out updated directory listings at these sites. "Browsing through terabytes" of text, of software code, and of other data, has become the operative principle.

A number of other distributed information retrieval systems have also come to prominence in recent times. A system called WWW ("World-Wide Web") supports Internet-wide hypertext. Try it, with its simplest access interface: telnet to info.cern.ch (no password). WAIS ("Wide Area Information Server") is a network-wide information retrieval tool. To quickly try out WAIS, the author indexed 125 papers and book chapters in one of his subdirectories. The indexes were established for approximately half a million words in a few minutes. A natural-language query (Fig. 5: "Everything you've got on the VLT . . .") yielded a thoroughly acceptable set of documents. This tool, though, aims at *distributed* information retrieval, accessing remote servers. If the information exists, and is set up approximately for WAIS, then it will be found. WAIS represents one line of attack on the "resource discovery" problem. It is a step on the way towards unleashing the genie in the computer networks which are now commonplace in research.

6. From Palette of Tools to Chiaroscuro of Knowledge

The developments overviewed here are driven by one consideration: there is no alternative. Mounting quantities (and

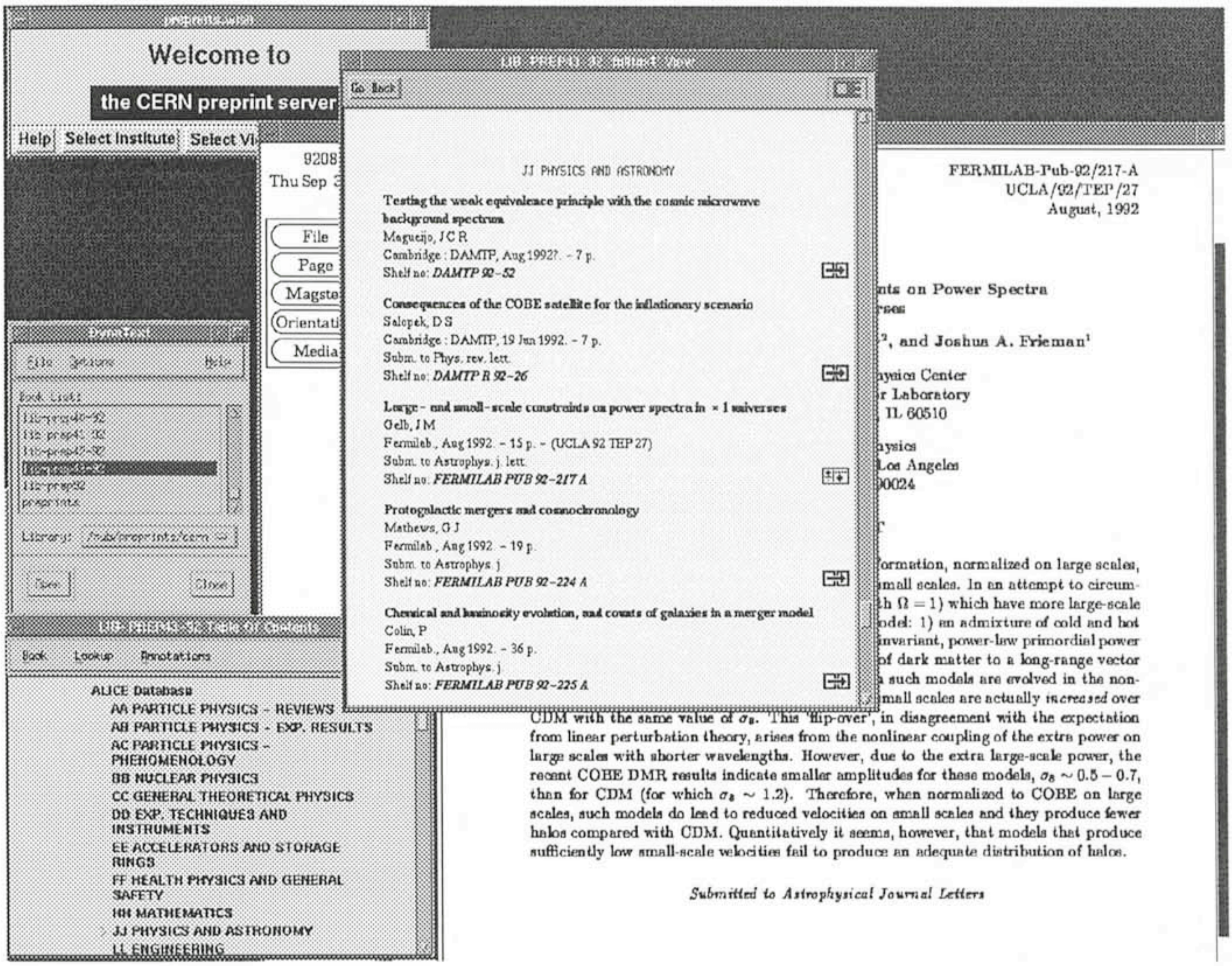

Figure 4: *The CERN preprint service. Shown are windows with a subject list, a recent preprint list, and the full text corresponding to one such preprint. A hypertext system facilitates navigation through all of this information.*

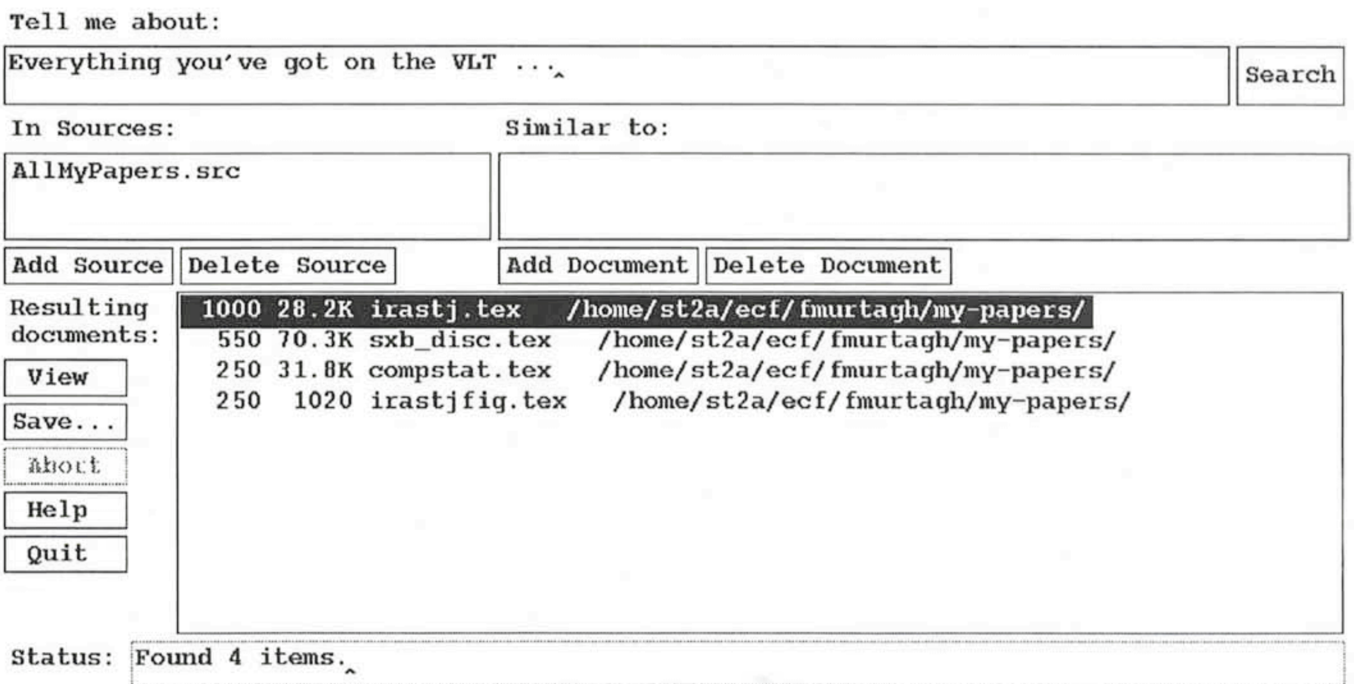

Figure 5: *A WAIS query ("Tell me about:") and results (on a set of the author's papers, which had been indexed beforehand).*

quality) of data require such steps to produce the global and integrated picture of information and knowledge.

7. To Probe Further

For material throughout, see inter alia Albrecht and Egret (1991), and Heck and Murtagh (1992).

Section 1: See Pirenne and Ochsenbein (1990 – to be updated soon).

Section 2: For STARCAT, see Pirenne et al. (1992).

Section 3: For ESIS, *set host* to *esis* (29671) and login with username *esis* (no password); or telnet to *esis.esrin.esa.it* (192.106.252.127), again using username *esis* with no password. On the Correlation Environment, see Giommi et al. (1992).

Section 4: For SIMBAD, see Egret et al. (1991). For contact points of commercial database providers, see Watson (1991). On the CERN preprint server, see van Herwijnen (1992).

Section 5: For FAQs, see Higgins and Leech (1992). For archie see Feigelson and Murtagh (1992). For archie, Gopher, WAIS and WWW, see contributions in Heck and Murtagh (1993).

Acknowledgements

Thanks to: M. Albrecht, R. Hook, B. Pirenne and R. Albrecht, for comments, and to S. Ansari and E. van Herwijnen for some of the material used.

References

1. M.A. Albrecht and D. Egret (Eds.) *Databases and On-Line Data in Astronomy*, Kluwer, Dordrecht, 1991.
2. D. Egret, M. Wenger and P. Dubois, "The SIMBAD astronomical database" in M.A. Albrecht and D. Egret, Eds., *Databases and On-Line Data in Astronomy*, Kluwer, Dordrecht, 1991, 79–88.
3. E.D. Feigelson and F. Murtagh, "Public software for the astronomer: an overwiew", *PASP*, **104**, 574–581, 1992.
4. P. Giommi et al., "The ESIS Correlation Environment prototype", in D.M. Worrall, C. Biemesderfer and J. Barnes, Eds., *Astronomical Data Analysis Software and Systems I*, Astronomical Society of the Pacific, San Francisco, 1992, 59–61.
5. A. Heck and F. Murtagh (Eds.), *Astronomy from Large Databases II*, ESO Conference and Workshop Proceedings, 1992.
6. A. Heck and F. Murtagh (Eds.), *Adding Intelligence to Information Retrieval: The Case of Astronomy and Related Space Sciences*, Kluwer, Dordrecht, 1993, forthcoming.
7. W. Higgins and J. Leech, "Compiling answers to frequently asked questions about space on computer networks", in *Proc. 43rd. Congress of the International Astronautical Federation*, Aug. 28–Sept. 5, 1992, in press.
8. B. Pirenne, M.A. Albrecht, D. Durand and S. Gaudet, "STARCAT: An Interface to Astronomical Databases", in A. Heck and F. Murtagh (Eds.), *Astronomy from Large Databases II*, ESO Conf. and Workshop Proc., 1992.
9. B. Pirenne and F. Ochsenbein, "New media for data distribution and exchange in astronomy", *ST-ECF Newsletter*, No. 12, Jan. 1990, 11–12.
10. E. van Herwijnen, "Are document servers replacing journals?", preprint AS/MI-92-05, CERN, 1992.
11. J.M. Watson, "Astronomical bibliography from commercial databases", in M.A. Albrecht and D. Egret, Eds., *Databases and On-Line Data in Astronomy*, Kluwer, Dordrecht, 1991, 199–210.

ESO Computer Networking

R. Hook, ST-ECF

1. Introduction

Astronomy has always been an international subject from the historical link to navigation up to the modern requirement to erect telescopes at the best sites in the world regardless of distance. ESO is itself a fine example of this trend – what could be more international than a collaborative organization of eight different countries running an observatory in the opposite hemisphere? Efficient operation requires efficient communications and in the era of predominantly digital data and text processing this means efficient computer networking.

Networking has now advanced enough that it reaches most astronomical sites world wide. Operations such as sending electronic mail from an Institute in Estonia to a student observing at La Silla are now taken for granted although they would have been unthinkable in more ways than one twenty years ago. Another important recent development, which is totally dependent on networks, is remote observing in Chile directly from an institute in an ESO Member State without either the astronomers having to travel to Garching to the remote control centre or all the way to Chile. This successful experiment was described in detail in the last edition of *The Messenger*.

Despite these huge improvements, networking still has a long way to go before it becomes as consistent and easy to use as the telephone or FAX machine. There are several different networks in use and they all have their quirks, foibles and inconsistencies. They are also often too slow and sometimes don't work. This article describes briefly the most important current networks used by astronomers. It then describes in more detail how external users may contact ESO electronically and what facilities are available. People who are already familiar with the networks may find most of the important information they need in the box which summarizes ESO electronic contact points. The text inevitably uses rather a lot of acronyms for conciseness. These are explained in Table 1 which should be consulted when necessary.

2. The Main Networks

Computer networks have tended to expand from modest systems linking workers in a similar discipline or geographical area to huge "internets" spanning the globe and united by the use of a common protocol. The protocol may be thought of as the standardized set of rules for communication which is independent of the type or manufacturer of computer equipment. There are now two main "protocols" which are dominating international science networking. They are the TCP/IP protocols used by the Internet and the set of standard protocols defined by the International Standards Organization (ISO) and often referred to as OSI. In addition there are several other protocols in use which astronomers encounter, in particular DECnet (used by the SPAN network), Bitnet/EARN and UUCP. To some extent these may be used together. For example it is quite common for TCP/IP or DECnet to be implemented "on top of" the lower level OSI protocol X.25.