

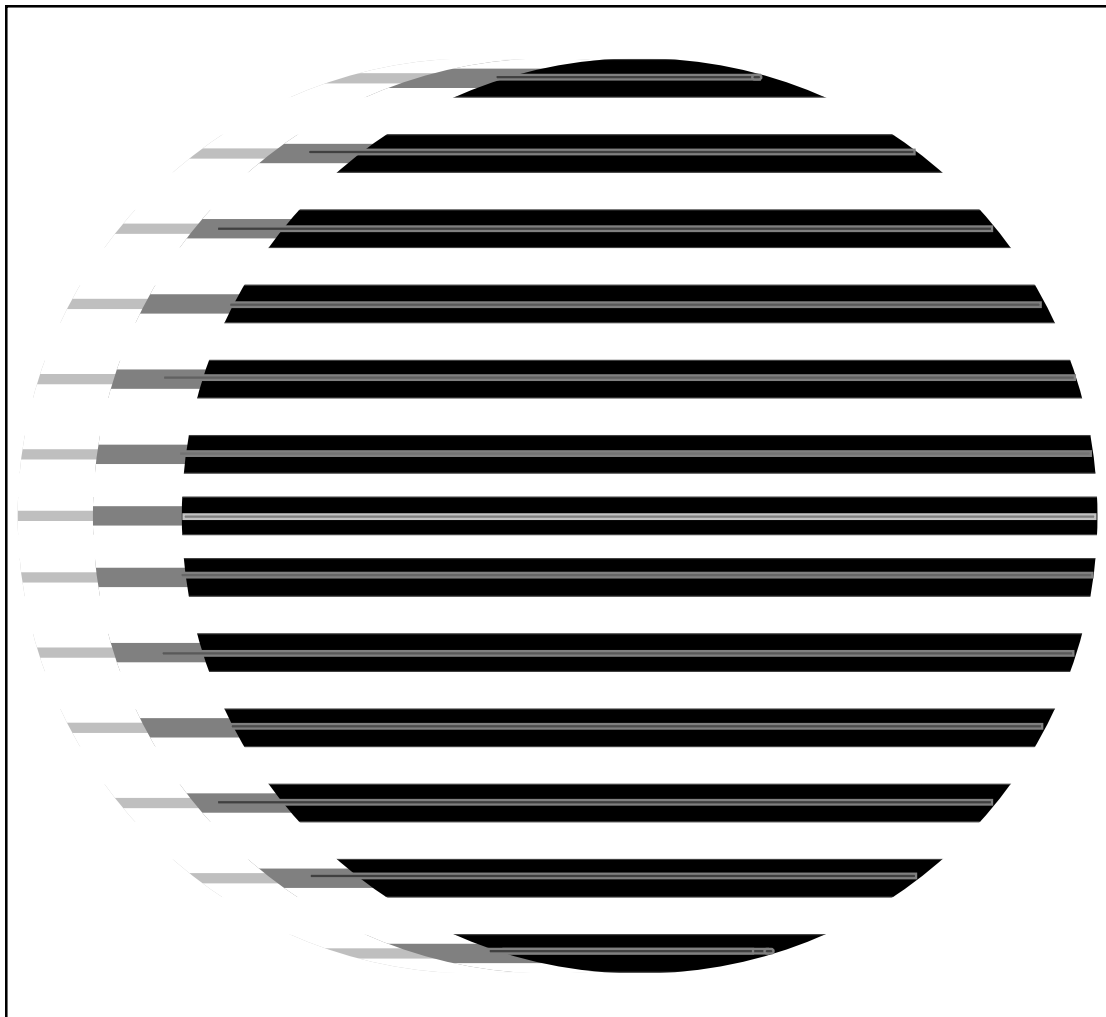
IASSIST

Q U A R T E R L Y

VOLUME 20

Summer 1996

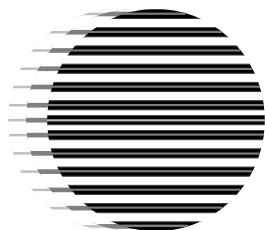
NUMBER 2



Printed in the U.S.A.

IASSIST

QUARTERLY



The **IASSIST QUARTERLY** represents an international cooperative effort on the part of individuals managing, operating, or using machine-readable data archives, data libraries, and data services. The **QUARTERLY** reports on activities related to the production, acquisition, preservation, processing, distribution, and use of machine-readable data carried out by its members and others in the international social science community. Your contributions and suggestions for topics of interest are welcomed. The views set forth by authors of articles contained in this publication are not necessarily those of **IASSIST**.

Information for Authors

The **QUARTERLY** is published four times per year. Articles and other information should be typewritten and double-spaced. Each page of the manuscript should be numbered. The first page should contain the article title, author's name, affiliation, address to which correspondence may be sent, and telephone number. Footnotes and bibliographic citations should be consistent in style, preferably following a standard authority such as the University of Chicago press *Manual of Style* or Kate L. Turabian's *Manual for Writers*. Where appropriate, machine-readable data files should be cited with bibliographic citations consistent in style with Dodd, Sue A. "Bibliographic references for numeric social science data files: suggested guidelines". *Journal of the American Society for Information Science* 30(2):77-82, March 1979. If the contribution is an announcement of a conference, training session, or the like, the text should include a mailing address and a telephone number for the director of the event or for the organization sponsoring the event. Book notices and reviews should not exceed two double-spaced pages. Deadlines for submitting articles are six weeks before publication. Manuscripts should be sent in duplicate to the Editor: Laura Bartolo, Libraries & Media Services, Kent State University, Kent, Ohio 44242. (216) 672-3024. Email: LBARTOLO@KENTVM.KENT.EDU. Book reviews should be submitted in duplicate to the Book Review Editor: Daniel Tsang, Main Library, University of California P.O. Box 19557, Irvine, California 92713 USA. (714) 856-4978 E-Mail: DTSANG@ORION.CF.UCL.ED

Title: **Newsletter - International Association for Social Science Information Service and Technology**
ISSN - United States: 0739-1137 Copyright 1985 by **IASSIST**. All rights reserved.

C O N T E N T S

Volume 20

Number 2

Summer 1996

FEATURES

- 4** **How to Make Osiris More Interoperable While Waiting for the Ideal System**
by Lennart Brantgarde & Leo Rubinstein ββ
- 6** **The Need to Train Librarians in Data Related Issues**
by Julia Dawn Paris
- 12** **Training in the Age of Digital Abundance - Technology or Information?**
by William Feigelman & Thomas Taylor
- 18** **Data Services in a Non-Academic Environment**
by Ineke A.L. Stoop

How to Make Osiris More Interoperable While Waiting for the Ideal System

by Lennart Brantgarde and Leo Rubinstein¹,

The concept of interoperability will in this paper be understood as a quality that makes a system flexible over time and adaptable in relation to other systems. It stands for high portability but also for high usability for different purposes.

The Osiris format was long ago sentenced to death. Its death-struggle has now been stretched into an extremely painful decennial process and I would guess that there will be several years more to come before it is over. The reason for this extended process is not so much dependent upon an abundance of virtues of Osiris but on the lack of alternatives. Osiris is amazingly enough still the only format that easily can reach a high interoperability versus a variety of functions that you have to operate at a modern dataoriented archive.

Two contrary principles for a data archive are:

1. Archive and store what you get and disseminate what you have
2. Transform what you get into a format that enables you to disseminate what is wanted in terms of formats by the market.

We suppose no one operates exactly on one of these extremes - we are located on the continuum in between, some more to the first position, some more to the second. This paper deals with how to achieve a position close to the second while still using Osiris as the basic archival format for preservation and documentation.

As you know one of the basics of Osiris is its oldfashioned fixed format. This makes it easy to program utilities for since every single unit of information always is found at the same spot. One way of making Osiris more interoperable is to create utilities. And we would guess that a whole set of utilities have emerged at various data libraries over the years but unfortunately they have never been collected and distributed to the benefit of all. At our site we have therefore been forced into this kind of utility production. We are going to present five such utilities and show how you can get hold of the programs.

From Osiris type 1 to Osiris type 3

In the first place we had to convert an Osiris-type-1-codebook into an Osiris-type-3-codebook. The binaries in the

type-1-codebook were of no use in the long run at least if you want full control of the file. We created a program that is called **123**.

This program comprises two things

1. Conversions of binary numbers to decimal codes.
2. Conversion of character codes, usually from EBCDIC to ASCII

As far as the binary numbers are concerned a complicating fact is that some data entries in a type-1 codebook are stored binary in 2 bytes and can have values between zero and 65336. The same information in a type 3 codebook has only four digits reserved and they are supposed to be stored decimally. Consequently no higher numbers than 9999 can be generated.

Our solution to this was the gv-convention. This comes into consideration only for numbers above 9999 and consists of a hexadecimal kind of system where zero thru nine and capital A thru capital F have been replaced by the minor letters g thru v where these letters stands for zero thru 15. In other words base 10 has been replaced by base 16 which allow us to take care of all binary numbers without losing any information and at the same time preserve the basic format of type 3.

The second problem - moving from character code representation to another - was solved by using, as a second parameter, a name of a file containing a translation matrix, a table with 256 lines. It is up to the user to define this table or modify existing ones. In such a table there are just two columns where the left one has entries equal to the line number -1 and the right one tells what character should replace the former. Value -1 in the right column means that the code is undefined. When encountering such undefined values the program shows the partially converted line and asks for a character to put in.

The program can be initiated by the command:

123 codebook [table]

where codebook is the stem of name of the codebook file and table the name of the translation table file. The program assumes that infile is codebook.os1 and writes the outfile codebook.os3. If the optional parameter table is left out the program assumes that the existing internal table named ascii

should be used.

Making an eyefriendly document out of an Osiris-type-3-codebook.

The original basic fixed formatted Osiris codebook file is not very useful for having as a readable document on the bookshelf. You have to get rid of all abundant figures and rearrange the information into a userfriendly textbook. This is done in our environment by a program called

kbl

The program is initiated by the command:

kbl infile outfile [dumpfile]

where `infile` defines your osiris-type-3-codebook-file and `outfile` defines what file you want to get out. In a unix environment the outfile can be browsed using the unix utility `more` and printed using `lpr`.

Moving an Osiris-type-3-codebook into World Wide Web or AUIS.

A modern Data library or a modern Archive has to consider desktop browsing of the archival holdings. This is absolutely necessary for an archive that has to deal with a scattered academic community located miles and hours from the office. To our help facilities like WWW, Netscape and Mosaic have emerged like benevolent fairies out of the dark.

Like the original fixedformatted Osiris the modern WWW is carrying an heritage from the childhood of computing. To get a text into good order in a web you have to tag it into HTML which is a sample of SGML. The next utilityprogram to be presented here transforms an Osiris-typ-3-codebook into an html-tagged document that, put out on the server, not only makes use of bold characters and italics but also automatically creates internal weblinks between the table of contents and the rest of the codebook document. In one stroke a stereotype Osiris codebook is turned into a decent modern eyefriendly document and a fullyfledged hypertext document.

The program is called `htkbl` and is started by the command

htkbl infile outfile [dumpfile]

exactly repeating the moments of the former program. `Dumpfile` is the name of a file to where wrong lines in a codebook (Infile) will be copied if any. For the AUIS-environment a utility program called `atkkbl` can be used the same way.

Getting from Osiris to SPSS and further out

So far we have just talked about utilities facilitating the infomative functions of an archive - to get Osiris ready for printing and make it exposable through modern networks. Fundamental to an archive is to disseminate both data AND documentation.

In order to facilitate distribution of data in a variety of formats we have created a utility that enable us to go automatically from Osiris-type-3 to a portable systems file in SPSS.

The program is called **otosp** and is started thru command

otosp codebook setup datafile exportfile

where `codebook` is the name of your Osiris codebookfile (Type 3), where `setup` is the name of file including all Spss control cards -including value labels -, where `datafile` is the name of your Osiris datafile and where `exportfile` is the name of your outfile created by SPSS.

The program can also be started by its mere name

otosp

in which case the program will be prompting questions for all other parameters.

Once in SPSS there are other conversion programs that can bring you further out into the djungle of statistical packages.

How do I get hold of these utilities?

All of them are stored under /pub/programs in

ftp://oden.ssd.gu.se

You are welcome to pick what you want. Source codes to 123, `otosp`, `kbl`, `atkkbl` and `htkbl` are free to be copied, compiled and used. But it is to be noted that these programs were created to suit the needs and requirements at the SSD and so far they have been tested only here. You should pay some attention to readme-files in subdirectories for improving your chances to install and customize the programs for your environment.

For those of you who would like to have a training example we have also made available a codebook of the oldfashioned ICPSR type, the codebook to Almond and Verba The Civic Culture. This exist in a type-1 binary version and is supposed to be converted to a typ-3 ascii version and further on to WWW and SPSS.

For installing copy a complete subdirectory respectively to your computer read the read-me file modify the make file: change the dir-string to the name of an appropriate directory on your computer (in search path for executable files). run the make file.

Further questions should be addressed to:

Leo.Rubinstein@ssd.gu.se

Good Luck!

1. This paper has been presented at the CSS96/IASSIST conference at Minneapolis, University of Minnesota, May 12 - 19, 1996. It describes a set of programs that have been built around a major documentation system called A-SIDE developed by Stephen Greene for Swedish Social Science Data Service and presented in Stephan Greene: A Functional approach to Documentation and Metadata, IASSIST Quaterly vol 19 no 1, spring 1995 pp18. Support for this presentation has been received from the Swedish Institute, Stockholm, grant 989/301/40

The Need to Train Librarians in Data Related Issues

by *Julia Dawn Paris*¹
University of the Western Cape

Introduction

The training of librarians in data related issues has become an educational imperative emanating from the technological advances experienced globally. The dramatic strides in technological developments are causing major expansion in the way that end-users access information. This sudden explosion of information has caused a shift from the print to the electronic mode, which in turn has led to a shift from ownership to access. Academic librarians make information accessible for research in academic libraries by placing at the individual's disposal vast resources of information to satisfy information needs. Academic librarians are thus challenged to redefine their service philosophy. A reconsideration of roles and mindsets are the profound changes that will have to be effected.

Parallel to the developments in technology, the political changes and the unfolding transformation process in South Africa necessitates academic librarians to reconsider their information provision strategies to redress imbalances.

Statements of Objectives

In this paper I will focus on the following:

- (i) the impact of technology on academic libraries;
- (ii) the impact of technology on the client of the academic library, with special reference to the University of the Western Cape; and
- (iii) the changing roles that academic librarians are required to play.

The abovementioned issues will be contextualized within the framework of the political changes occurring in South Africa.

Clarification of Concepts

For the purpose of this paper, data is defined as facts, statistics or information that can be analysed. Knowledge is defined as familiarity gained by experience. It is also viewed to be a person's range of information or theoretical or practical understanding. Information is viewed as desired items of knowledge. One notes a diversity in the definition of information. This is in view of the fact that it is deemed to be intangible and only encountered operationally by its effects. It is suggested to be derived from data, and can be accessed through print and computerized machinery for utilization (Zorckoczy, 1988: 11).

Training as used in this paper refers to the concern with making the best use of the Human Resources in an

organization by providing them with the appropriate instruction to acquire the necessary skills for their jobs (Statt, 1991: 154).

Impact of Technology on Academic Libraries

Internationally, the age of computers have certainly impacted on the traditional manner of doing things in a library. The dynamic nature of information generation, management and use, as well as the proliferation of publications, force the library environment to either adapt or die. Microcomputers have streamlined the operational workflow of routine functions as well as enhance the online search process. Cataloguing and circulation automation are providing more effective service and better control over collections. Computer technology effected various ways of accessing information more speedier through networked electronic mail facilities. It now allows librarians to automatically logon to local and remote systems and download search results for later printing. This is causing a major shift from libraries' provision of information from own collections to that of access to remote regional, national and international collections, databases and networks. Advanced technology also gave rise to the virtual library - a concept used to denote remote access to the contents and services of libraries and other information resources, combining an on-site collection of current and heavily used material in both print and electronic form with an electronic network which provides access to, and delivery from, external worldwide library and commercial information resources. This is a transformation with dramatic results for research which cannot be ignored. Computer technology has indeed made it possible for libraries to establish networks, based on co-operation and resource sharing, with each other and with other information centres. The concept of networking, understood to be the building of contacts among professionals, has been given new meaning by the academic library's utilisation of technology. Here one thinks especially of the INTERNET, which is seen as the network of networks. It brings together people, databases, and networks. Through technology eager use is also made of periodical indexes and reference works on CD-ROMs (Compact disc read-only-memory). It is interesting to note that this transformation, brought about by technology, does merit a lot of adjustment. This is demonstrated in library school curricula which, today, include a module on Information Technology or a complete course on computer science. It is true to say that the increasing complexity and sophistication of information technology requires a high degree of specialized technical

knowledge. This requirement has far reaching implications for the future training of academic librarians. Academic librarians should learn how to exploit new technology to the benefit of their user's information needs. They have a responsibility to require the knowledge and skills necessary to use and teach the most efficient information techniques the current technology makes possible. A proactive and flexible manner is also required if academic librarians want to take up the challenge of technology, to interact and communicate effectively with clients who may not only be unfamiliar, but also uncomfortable with information technology. Therefore, training and retraining should be high on the strategic priority list of academic libraries.

In South Africa, academic librarians and end-users experience the convenience of information made easier and speedier by sophisticated machinery. On the other hand, they experience difficulty in the accessing of the information and information overload. Information overload occurs when locating too much information on the given topic. This dilemma creates a feeling of being overwhelmed and overloaded and this could result in a situation of frustration and anxiety. If a user is not able to understand the automated system used for an online catalogue at any academic library, or the systems potential to give information, or know how to sift and sort through the plethora of online information, it immediately results in information poverty by creating a barrier between the user and the information needed. Hence, information technology becomes a problem when it deprives users of information, especially if the information is to satisfy basic needs as experienced by the majority of people in South Africa.

South Africa is a 'new kid on the block' in the development of technology and the application thereof. Estimations claim that information technology is developing at an annual rate of 35% in South Africa. This is viewed to be very slow in comparison with developed countries where information technology is doubling, if not tripling, at that rate every four years. My major concern is with the extent of the impact of this development of technology, however slow, on the information impoverished people of South Africa. The South African academic libraries, and in particular the University of the Western Cape (UWC) have been influenced by the global information explosion. The access of our users to local, national, international databases and networks, has made it necessary for the librarians to take an objective look to what is happening. I agree with Makhubela (1995: 15) who states that "though this situation is not necessarily unique to UWC, it becomes especially critical, given that many students at UWC come from economically deprived and disadvantaged communities". The realization then, that technology is affecting the way that librarians provide access to information should become a catalyst for change. However, outdated mindsets and curricula make it downright impossible to take up the challenges posed by the shift and increased user expectations.

Challenges facing librarians in South Africa are not only caused by the technological developments, but also by the major political change from an oligarchic, Apartheid society to that of a democratic one. Great educational inequalities, major illiteracy problems, lack of a reading culture, lack of school libraries are some of the stark realities academic librarians are faced with. The Nationalist Apartheid government fostered a library and information service characterized by a traditional approach as was followed by many of the developed countries. It was focussed to benefit the educated user community, was literacy based, with collections comprising of predominantly books, and an emphasis on facilities and collections instead of the needs of users. In a nutshell it served the needs of the dominant white culture and class (Nepi, 1992: 54). The transformation process now underway calls for the library and information service to also undergo transformation and address the social responsibility of libraries. It calls for a shift in emphasis from collections to needs of users. It also challenges librarians to use information as a source of development and empowerment in the education of all those served. In the academic environment librarians need to learn to facilitate the learning environment of disadvantaged students even if it means teaching them step-by-step how to use the technology in its basic forms e.g. OPACS, and later teach them how to access CD-ROM databases and other remote online networks through Information Literacy Programmes.

The Impact of Technology on Clients

Industrial progress requires a versatile and skilled labour force. This could be effected by well planned education and training strategies. Information needs vary according to levels of education and socio-economic status. It is important to assess what it is that information end-users need and want. To end-users accessibility of information is one of the primary issues of concern in satisfying information needs. No matter how organized the information, it will not realize its value until it is made known and put to effective use.

A generalization reflected in international literature is the idea that undergraduates usually labour under time constraints. Also, that their information needs are most often met by the collections of their own institution. Further, that depth of reference queries are not so complex and that less reference assistance is required. This argument definitely assume a situation where undergraduates have been exposed to adequate education, school libraries, and have a sound economic background.

At an institution such as UWC, a lack of factors noted above, combined with problems of language, render undergraduates incapable of having a basic understanding of their research topic under investigation. Minimum or no knowledge of basic reference sources, inability to verbalise information needs are other shortcomings encountered by academic librarians.

Post-graduate and faculty needs are viewed as extending beyond the resources of their own institution. The research of these clients represent unique contributions to the knowledge base of disciplines. Therefore, their demands on reference services are more specific and exhaustive. For them the academic library extends into the resources of other information centres through inter-library lending services. With the advent of online searching on remote databases and networks 'the world become their campus'. The appearance of computer accessible information databases have changed the way researchers search for information. Some scholars rely almost entirely on computerized sources for current citations to published research in progress. This is creating great expectations toward academic libraries to provide the necessary mechanisms to access these sources.

At UWC our post-graduate and faculty needs do not differ as widely as that experienced internationally. However, with post-graduate students the librarians still have the responsibility to teach the most efficient information techniques the current technology in the library makes possible. I wish to reinforce Benson's (1995: 57-69) utterance when he said that "our enthusiasm about the new technologies should not outstrip our ability to provide adequate service to those for whom we acquire the services and may cause us to overlook real needs". I firmly believe that we have a responsibility to be sensitive to research needs of the entire academic community we serve. To reflect this sensitivity, we need to conduct use and needs studies to ensure acquisition of the databases and networks.

Changing Roles of Academic Librarians

Traditional role of librarians

Previously, librarians emphasised preservation of books, making them accessible through cataloguing and classification, and offering a client advice and guidance service. Librarians were only required to know where information could be found rapidly to answer the client's need or questions. Librarians were viewed as custodians of collections of library material instead of proactive individuals. This situation is still prevalent in many academic libraries in developing countries.

New roles of librarians

In the developed countries, the paradigm shift from traditional to electronic libraries have changed the way librarians execute their duties today. The expectations created by this paradigm shift resulted in the need for librarians to adopt a new vision and a new mission to meet users information needs. They are faced with new issues, and new challenges brought about by the increasing use of information technology in the library environment and the rapid changes occurring in this area, understanding of what librarians duties are and the new attitudes that should result from this change in mindset.

There is also the added challenge of helping clients to

manage the information overload that came about as a direct result of this new era. Clients must be assisted to sort through the wealth of information and make wise decisions about what they need. There is a challenge posed at academic libraries to shift the emphasis from physical documents to individuals with needs; from document delivery to information management and transfer; and from question-answering to problem-solving. Librarians are challenged to move beyond quantity to quality. A general view seems to be that although this shift has occurred, librarians continue to be more concerned with delivery of documents and have not started to focus on the delivery of contents or the data and information contained in the documents. My contention is that this is due to the failure to grasp that there is a need for this new vision. There should be a move towards acceptance of the facts that major paradigm shifts are creating new definitions of what librarians roles are to be in the technological library environment. We face expanded demands on our time and skills and are afforded a tremendous opportunity to assist students and faculty to solve their research problems. Requisite skills that need to be acquired include: competencies ranging from communication skills to knowledge of database searching techniques; sifting and analysis of information; and the reduction of the amount of information provided to the user. Larry Benson (1995: 58) postulates that "as attitudes toward the use of automated systems change, so must the roles of academic librarians. They face expanded demand on their time and skills and have a tremendous opportunity to assist students and faculty to solve their research problems".

The academic librarian is called upon to serve in overlapping roles namely as information specialist; teacher/trainer; and as consultant. Depending on the library's priorities, elements of all three roles will enter into the professional librarian's position. As information specialist, the librarian will be called upon to provide adequate access through adequate resources for optimum usage. As teacher, the librarian will have to teach library, information, and technology literacy skills. This includes the teaching of critical thinking skills to assist clients to become active, independent and confident users of information technology. As consultant the librarian is called upon to participate in academic curriculum design and assessment projects and to provide expertise in the selection, evaluation and use of materials and emerging technologies for the delivery of information and instruction, as well as translating curriculum needs into academic library program goals and objectives. A reluctance to assume this role would deny the faculty and administration the benefit of the valuable insights academic librarians have in these areas. It would also minimize the importance of the academic library to academic outcomes. This expanded role of the academic librarian, brought about by information technology has great potential for improved educational outcomes.

At UWC, the role played by academic librarians has not

changed very much. The new issues and challenges of technology have not resulted in a radical change of roles yet. In fact, the technology applied in information provision and access have not been exploited to its fullest potential. Lack of resources have made it impossible to initiate a process of exposing end-users to networks like the Internet. Language problems and other cultural barriers add to the existing complexities of users social backgrounds. The predominant roles thus played are primarily those of instructor in library skills, information provider, faculty liaison and collection developer. Lack of adequate staffing and specialized skills contribute to a work overload which are some of the factors slowing down the move toward the real role of facilitation and empowerment.

Recommendations for Training Strategies:

The training of librarians should be understood within the broad framework of Human Resource Development and Capacity Building within the transformation paradigm of South Africa. In order to play a proactive role in facilitating information access and its proper use, it is imperative to look at the role librarians can and should play. The suggestions posed in this section are not proposed as solutions, but should serve as useful information that could assist in appropriate applications of training strategies.

Any training or retraining embarked upon by librarians in South Africa, should be done in line with the needs of those wanting and needing the services. The bottom line for training should always be improvement and efficiency of services to clients. This is especially seen against the background of cultural diversity within the population as well as their previous exposure or non-exposure to information. Reference librarians need to be equipped to become more sensitive to cultural diversity and in this new technological environment develop more effective communication skills to understand the multi-cultural aspects of information seeking behaviour. In order to make a difference, these aspects should be inculcated through the curriculum design at Library school entrance level. Continuous education efforts to ensure lifelong learning in multi-cultural information provision and the evaluation of performance should also become a priority.

The majority of users of the historically marginalised institutions come from a background of poverty, illiterate parents, overcrowded homes, lack of proper housing, lack of reading facilities, oral tradition and lack of proper schooling, to mention only a few. For these users information is geared towards coping to survive. Librarians need to be equipped to understand this dilemma experienced by the majority of users and to devise strategies to effectively address it so that they become instrumental in helping users.

Information technology and other mechanisms should be applied to enhance the development process of information seekers in South Africa, not impede it. Librarians must be

equipped to break down the barriers which hinder information access. The unfolding transformation process brought about a new dimension of complexity in the form of all the challenges facing academic librarians. Librarians should provide an unthreatening atmosphere in which users could access the necessary information. Librarians should also become instrumental in computer aided instruction programme development.

Training must include elements of basic data analysis, networks, and network search commands, the nature of interfaces, search software of specific social science and other databases. General training in the nature of all databases and networks is imperative. Training in end-user instruction, to ensure optimum utilization of these costly resources, is necessary. In order to instruct end-users, librarians themselves need to be trained in how to access, search and interpret different databases.

The intention with strategies mentioned should be to use existing resources. The challenge is placed before tertiary institutions and other academic alliances to respond to the training and continuing education needs of the academic librarians by transforming themselves in the way they operate. This statement is made in view of the fact that there is considerable competence and skills existing within the various alliances of the Library and Information Service, but outdated principles regard users as passive consumers of information instead of an active participant in the information process. Within the South African situation, there are no simplistic answers. Here academic librarians face formidable challenges, but also lots of opportunities to make a difference.

Training should also address all levels and needs in terms of how to supplement current awareness of Information technology, how to do retrospective searches and how to use technology in the social science and other related fields.

Basic training is required to increase the awareness of people about the possibilities offered by the new technology and teaching them effective ways to make use of the equipment. This is the equivalent of teaching people how to use the computerized catalogues. A second kind would be training needed to become fully familiar with a complex word processing or database software package. If the training is neglected, the full benefits of the facilities will not be realized and the drawbacks will be magnified (Boston,G,1994:331-337).

The introduction of new electronic services has implications for virtually all library staff. Inevitably, as electronic information systems expand, workloads will change and resources, both human and financial, will need to be redirected. Successful implementation and user satisfaction in the electronic environment, will only be achieved by a highly trained and skilled library workforce. These will

include increased knowledge of automated systems and electronic communications, as well as the hardware such as workstation networks.

Furthermore, it is important that information workers continue to be active in influencing the newly emerging national and international standards for electronic information, e.g., standards for bibliographic control and for citing articles. "Some electronic network travellers will fearlessly set off on their own. Others will still rely on librarians to drive the tour bus" (Woodward, 1994: 44). In the UK, the Library Association is encouraging librarians and information workers to analyze the contents of online information systems, databases and CD-ROMs in the same way that librarians have critically analyzed the contents of reference books for accuracy and up-to-dateness.

The information era, as previously noted, puts before us more information than can be consumed and used. The sifting and analysis of information, and the reduction of the amount provided to the user, is an equally important responsibility that is based upon the traditional skills of information workers. The quality and standards of service provision has become an important factor too. Agreed and published standards or guidelines are mechanisms to be used, because they are important factors toward improvement in the provision of services. The professional librarian also has the responsibility to maintain his/her level of knowledge, skills and expertise given the speed of change enhanced by the information technology and telecommunications developments. It is our responsibility as well as those of employers to ensure that the knowledge base and expertise of information workers is kept up-to-date. It should be made mandatory by Library Associations.

Literature suggest a code of ethics or conduct as the machinery to exert pressure, should information workers fail to meet the responsibilities as set out above. In training, librarians should be charged with learning the structure of online libraries, the mechanics of searching, and the relationship between database content and user questions. Librarians should be trained not only to focus on specific databases within their areas of specialization, but also to look at general news databases which transcend subject discipline. To understand this magnitude of online library collections, a major conceptual shift in thinking must occur. The ability to search full-text articles makes the service unique, and poses special challenges. The importance of constructing precise search statements, critically selecting words and utilizing appropriate commands to narrow the retrieval are some of the special instructional challenges for reference librarians (Adalian & Rockman, 1995: 99-113).

Identifying, locating and using networked information can only be accurate with the help of an intermediary - specialized information workers who are the links between the information seekers and the information in the databases.

In the networked environment it is easy to become suddenly overwhelmed with a large volume of information. In the absence of user driven filtering of information librarians are called upon to play the role of filtering and interpreting the information for those who choose to use libraries.

Stakeholders to Assist in Training

The many challenges posed to academic librarians cannot be carried by them alone. What will be required, are the effective partnerships of all major stakeholders within the field of Library and Information Services. Training structures for librarians already exist, but should be effectively harnessed for continuing lifelong education. Stakeholders who should play a vital role in the education and re-education of academic librarians in South Africa are the African Library Organization (ALASA), the Inter Resource Forums, The Library and Information Workers' Organization (LIWO) and the South African Institute of Librarianship. These organisations should pool their resources and professional skills to provide a vehicle for Information Workers throughout South Africa to benefit from a broad range of professional expertise.

Other stakeholders identified are: the vendors of technological and commercial products - by giving basic training in their software; library schools - by updating and redesigning curricula as the changes in the information provision environment occur; University and Academic Library Administrators - through strategic planning of inservice training programmes in word processing, database-, online network searching and information management skills; and finally Government Departments who should make the necessary funds available to acquire the technology and training facilities needed to empower and build the capacity of academic librarians.

Conclusion

The training of academic librarians in data related issues is crucial if they want to play the new role of facilitator in the technological environment of the academe. To be able to do so academic librarians have to leave their comfort zones as custodians and become proactive role players in the information provision for development and progress. It is my belief that South Africa can learn valuable lessons from the international experience. However, priorities for training need to be constructed in accordance with the prevailing needs and conditions in South Africa, and not simply in relation to abstract theories or first world research findings.

References

- Adalian, Paul T & Rockman, Ilene F. (1995) Issues in implementing and teaching the Lexis/Nexus services in an undergraduate library. *The Reference Librarian*, 48: 99-113.
- Benson, Larry D. (1995) Scholarly research and reference service in the automated environment. *The Reference*

- Librarian, 48: 57-69.
- Boon, J.A. (1992) Information and development towards an understanding of the relationship. *South African Journal of Library and Information Science*, 60 (2): 63-74.
- Boston, George (1994) New Technology - Friend or Foe. *IFLA Journal*, 20 (3): 331-337.
- Bowden, Russel (1994) Professional Responsibilities of Librarians and information Workers. *IFLA Journal*, 20 (2): 120-129.
- Chatman, Elfreda A. & Pendleton, Victoria (1995) Knowledge gap, information - seeking and the poor. *The Reference Librarian*, 49/50: 135-145.
- Diaz, Karen R. (1994) Getting started on the Net. *The Reference Librarian*, 41/42: 3-24.
- Faries, Cindy (1994) Reference librarians in the Information Age: Learning from the past to control the future. *The Reference Librarian*, 43: 19-28.
- Figueiredo, Nice (1992) Information as a tool for development. *The International Information and Library Review*, 24 (3) Sept.: 189-201.
- Gerber, P.D; Nel, P.S. & Van Dyk, P.S. (1994) *Human Resource Management*. Halfway House: Southern Books Publishers.
- Hopkins, Richard L.(1995) Countering information overload: the role of the librarian. *The Reference Librarian*, 49/50: 305-333.
- In Focus (1995) *Human Science Research Council Bulletin*, Feb/March: 37-43.
- Library and Information Services. (1992) *Report of the NEPI Library and Information Services Research Group: a project of the National Education Co-ordinating Committee Cape Town: Oxford University Press.*
- Liu, Mengxiong. (1995) Ethnicity and information seeking. *The Reference Librarian*, 49/50: 123-133.
- Makhubela, L & Koen Z. (1995) Another angle on access: information literacy and student Learning. *Academic Development*, 1 (1): 13-19.
- McCallum, Sally H. (1994) Connectivity for Libraries and Information Services - Introduction. *IFLA Journal*, 20 (2): 145-146.
- Mchombu, Kingo. (1991) Which Way African Librarianship. *International Library Review*, 23: 183-200.
- Pask, Judith M & Snow, Carl E. (1995) Undergraduate Instruction and the Internet. *Library Trends*, 44 (2) Fall: 306-317.
- Sisson, Loren & Pontau, Donna (1995) The changing instructional paradigm and emerging technologies: New opportunities for reference librarians and educators. *The Reference Librarian*, 49/50: 205-215.
- Statt, David (1991) *The Concise Dictionary of Management*. London: Routledge.
- Summerhill, Craig A. (1994) Connectivity and Navigation: An overview of the Global Inter-Networked Information Infrastructure. *IFLA Journal*, 20 (2): 147-157.
- Sykes, J.B. (Ed.) (1982) *Concise Oxford Dictionary*. Oxford: Clarendon Press.
- Weibel, Stuart R. (1995) The World Wide Web and emerging Internet Resource discovery Standards for scholarly literature. *Library Trends*, 43 (4) Spring: 627-644.
- Westerman, Mary. (1989) Computers in Academic Libraries. *Health Science Libraries - micros and medicines. Catholic Library World*, 60 (4) Jan/Feb: 150.
- Woodward, Hazel. (1994) The impact of Electronic Information on Serials Collection Management. *IFLA Journal*, 20 (1): 35-45.
- Zorckoczy, P. (1988) *Information technology*. Oxford: Clarendon Press.
1. This paper has been presented at the CSS96/IASSIST conference at Minneapolis, University of Minnesota, May 12 - 19, 1996.
- Julian Dawn Paris, Subject Librarian - Economic and Management Science, University of the Western Cape, Private Bag X 17, Bellville, South Africa.

Training in the age of digital abundance - technology or information?

by Nicky Ferguson and Lesly Huxley¹,
University of Bristol

Keywords: Internet, training, information skills, World Wide Web

Information. Technology. Networks. Training. Four words which are often thrown together and thrown around quite carelessly. Our view is that in whatever combination they occur, technology, to quote Janis Joplin, always seems to come out on the top. Our view is that the present and the future demand that we take training, and training in information skills in particular, very seriously. Only then can we deal with the challenges which the advances in networked technology presents. Only then can we begin to help others to come to grips with new possibilities and use existing skills and experience to avoid new versions of old mistakes.

Nicky Ferguson writes:

Let's start where everyone should start these days - on the Web. Often people have noticeboards in their kitchens - I find these quite compulsive reading. They are covered with the essential detritus of individual, or family, life. Business cards from the plumber and the piano teacher; appointment cards from the dentist, the clinic and the acupuncturist; receipts from the washing machine repair woman and the milkman; shopping lists, opening hours, bus timetables, parking tickets. Sometimes I learn something from these unauthorised perusals ("gosh the ante-natal clinic! congratulations") but mostly the charm resides in glimpsing the ordinariness, the minutiae of other people's lives. Nosiness, not to put too fine a point on it. I'm not sure that trawling through most Web pages is very different. A superficial fascination but often no new information, no new ideas. This is actually fine - we don't advise people to keep their noticeboards covered with a black cloth in case someone else wastes their precious time reading personal trivia. It is up to me to discipline myself not to spend all day browsing the appointment cards. We should be explaining this to children, students, trainees, users call them what you will. It is not the Web's fault that people find it useful for collecting and collating their personal and work trivia and signposts, and sharing that information with their friends. It is up to the browsers, and here I mean the people not the software, to recognise quality when they see it and when they don't, and to develop strategies to make their work time more productive: and of course it is up to the professional providers of quality information to point people to worthwhile sources and to run quality services.

So what are these quality services? How will our students recognise them? Equally important how will they recognise those that aren't worth spending their time on? Moreover, since the line between information consumers and information providers will become increasingly blurred, how can we encourage them to make their information available in a neighbourly, useful, socially responsible, creative, even fun way? Let's look at some sites and see what we think of them.

Let's say I'm interested in goldfish - someone's told me that I should look at "**Sharon's home page**", OK, let's explore.

H' mm- it seems Sharon likes goldfish too but there's not really much more about goldfish than that. There seems to be a Jane Austen archive and there are other things which may or may not interest us, but no real material about fish; still there is a list of other places to look for fishy stuff including Dave's page - "**really interesting**" it says - so we'll go and look there.

Dave's page is strong on hyperbole but again it lacks content. Still it does have a list of fish sources (which look similar to the ones on Sharon's page, now you come to mention it). There seems to be a heavy metal music archive and there are other things which may or may not interest us, but no real material about fish; still there is a link to the "**SOFA home page**" (I know you've all heard of the Small Orange Fish Association) so we'll go and have a look there.

Well this isn't quite what I expected - this seems to be the home page for my kinswoman Finlay Ferguson, rabid Scot and fish fan, as well as chairwoman of SOFA. It does have a list of fish sources (which look similar to the ones on Sharon's page, now you come to mention it). There seems to be a clans and tartans archive and there are other things which may or may not interest us, but no real material about fish; still there is a link which reads "**Goldfish lovers click Here**" - now that sounds exactly what we're after, so we'll go and have a look there.

Oh dear... "**Sharon's home page**" ... back where we started.

One could argue that the act of classification itself ("Some fish pages I have gathered together") adds value to the web; but in fact a classification which only points at further pointers obfuscates rather than clarifies. Which is all a roundabout way of saying that many web pages seem to be a part of a self-referential, charmed (but not charming) circle,

merely pointing to each other without adding much to the sum of human knowledge or even networked information. Of course a classification which also describes, not a mere listing but a descriptive record, meta-data in the jargon, is a different matter. Depending of course on its own provenance and reliability, such meta-data does add value to the web and to the resources it describes and it can be amply justified. Care should also be taken to point either to resources themselves or, in some cases, to further, fuller or more specialised descriptive lists, but not to bare lists of titles, whether or not they are long or have hyperbolic introductions.

What implications does this have for us when we attempt to train others to construct worthwhile web-based resources? What should the resources do?

1. **Value**

They should add value in some way, probably providing descriptions of the resources they point to, preferably doing more.

2. **Classification**

They should systematically categorise and take advantage of the uniquely “virtual” nature of their medium to cross-classify, so that users can get used to knowing where to look and can find things where *they* expect to not where *we* think they should.

3. **Maintenance**

Resources should be continuously maintained to give currency in addition to reliability; networked information changes so fast that unmaintained lists go off quicker than milk on the doorstep.

4. **Quality**

They should not just uncritically dump everything that might be relevant into a huge list. Quality judgements should be made and continue to be made so that resources which are set up in a burst of enthusiasm and left to wither and become irrelevant are spotted and deleted.

5. **Variety of access**

Users should be offered a variety of access methods or interfaces - they should have the choice of searching or browsing and preferably have different browsing options.

And what are the implications for the searchers, the users (and those who seek to teach them)? There will probably be no “one way” of doing things and no one source or megastore which will satisfy all your information needs. You may expect to call at three or more locations and use different techniques before arriving at your goal. You should take care that you are narrowing the field along the way, not skipping from one unordered list to another. Reward the productive search paths by taking a few moments to retrace your steps and add bookmarks, penalise the not-so-charming

circles by making a mental note to avoid them in the future. So the Tao of webbing will be that there are many paths to enlightenment, you will browse and then search, search and then browse. In searching, how to search? In my experience the information strategies of the average user are limited. Many if not most of the postgraduate students I come into contact with have managed to get good degrees without knowing what the three letter word “and” means when it is used as a Boolean operator. They will search for “Marx and Engels” (or often “Marx and Spencer” but here is not the place to consider the quality of secondary education, spelling or the commodification of culture) and expect, very reasonably if no-one has told them otherwise, that they will find every resource which mentions “Marx” AND also everything containing “Engels”. Yet many sites providing search facilities on the Web will offer far more sophisticated options which go largely unnoticed or unused. There are two approaches to this problem. The first is the one that the technologists adopt. What we need to do, they tell us, is to make the search engines, the software, the facilities, so clever that users don’t need to know about search strategies, they can just type in their queries in natural language. I’m a big fan of natural language searching, at least I will be when I find a system that works, but I doubt if most users have given sufficient thought to what they are seeking, even to phrase their quest in natural language. Perhaps I will be convinced by someone here today who has designed a natural language search engine with artificial intelligence, in-built dictionary/thesaurus and pre-search feedback mechanisms so that when a user enters on the search form the word “aids”, before searching the database or sending the robot off to examine the web, our intelligent engine will ask the user “Do you mean handy gadgets for disabled people to allow them to operate machinery, pick things up, hear better and things like that, or do you mean the disease or do you mean home helps or do you mean something else I haven’t mentioned here?”. Even in that unlikely event I will still maintain that a sophisticated approach to searching will encourage sophisticated thought and that surely sophisticated thought is needed for sophisticated analysis. Of course I do not think we should discourage the development of excellent search mechanisms - I am in fact involved in a project in which we spend a lot of time discussing exactly what such a mechanism should and should not do for the user - but alongside the technological development, we should be encouraging and promoting amongst so-called “ordinary users” an understanding of strategies for finding, retrieving and using information.

Which brings us to training. You will have guessed by now that I think training should encompass more than the latest technological buzz, more than which buttons to press in version *x* of software *y* which will be replaced by software *z* in a few years, months or weeks. We are after some more general appreciation of ways to use these technologies for real work, even real life. I used to think that it was important to make Internet training a totally pleasant and stress free

experience, newcomers tend to bring quite enough stress and anxiety with them when attending a course on such a daunting and overhyped subject. I now think that it is somewhat mischievous and misleading to prepare and engineer such things to gloss over the difficulties and make everything too smooth and easy. Training, when related to preparing yourself for other activities such as running a marathon, swimming the channel or even a walking holiday, implies a certain amount of effort, dedication, commitment and practise. You make yourself, or your trainer makes you, do unpleasant things, push yourself, stretch, extend your capacity - you will be subjected to nauseating exhortations such as "no pain, no gain". Perhaps we should be taking the "make 'em sweat" approach a bit more in this area too. Of course it is necessary to present beginners with step by step practical exercises outlining every key press, what are known as hand-holding exercises. But if that is **all** we do, however impressive our evaluation sheets at the end of the day, we are not giving them the confidence to go further on their own. Better to suffer a few adverse comments at the end of the training day but produce trainees who, when the hand is taken away will wobble off on their own bicycles and disappear round the corner, not collapse in a heap. So let's make our poor trainees answer questions, don't just let them follow the instructions, or wander off on their own. We should, of course, provide reference materials and the equivalent of reading lists, citations catalogues and bookshelves, but as many university lecturers have found, it's not always best to dole out photocopies as it can encourage the belief common amongst students that the act of clipping a photocopy into a ring binder osmotically transfers information and comprehension of it to the brain. Often better to force the unwilling student to search for and actually read the article before deciding whether it is worth copying and archiving in empty cornflake boxes. Similarly with exploring the networks - next time they will be able to cope better if this time they had to work it out from a sketch map rather than being led by the nose.

In summary, I would like to beg, plead and cajole you, as information professionals to share your skills with the horde naive users like myself who are blundering and about to blunder into this huge global virtual library that is creating itself. I am asking you to consider, going **OUT** of your institutions and talking to the people who are and will be using the Internet. Get involved with training initiatives and patiently explain that people have thought about information issues before Netscape was installed on their PC. Go to the places where they are beginning to use this stuff - the school classrooms, the undergraduate Internet clubs, the cyber-cafes and worse (yes it's a dirty job but someone's got to do it). Spread the word about information handling skills, information seeking skills and user-friendly information provision. Don't let the code-writers monopolise the new image of international networked information - they will reinvent the wheel if you let them and it will be triangular (but with retractable spokes and flashing lights). What's

more, as the provision and use of networked information explodes, the technology will change at least every couple of years. But information skills will become more relevant, more important, more marketable. Be there, or be triangular.

Lesly Huxley writes:

The task, when I was appointed SOSIG Documentation and Training Officer in mid-1995 was twofold: to produce SOSIG promotional, publicity and reference materials drawing attention to a service which had done 'some of the hard work' in searching for quality and relevant social science resources; most academics who tried the Web when it first emerged from their University computer experts' clutches found it a significant time-waster and severely wanting. We wanted to bring them back into the fold, bringing the newcomers with them, to show them that there were ways of locating useful networked information quickly. Secondly I was to provide Internet workshops at UK universities and colleges of Higher Education, supported by training materials tailored for social scientists and the particular needs of the site concerned. The target audience comprised mainly newcomers to networking in the Social Science field and those tasked with training and supporting them. The workshops and materials were not to be set entirely in the 'press that button' mould - although newcomers would need some precise instruction, the aim was to provide a forum for learning both the tools and techniques *and* an attitude of enquiry which would allow them to cope with and extract the best and most relevant information for their work not only on the day of the workshop but well into the Net future - a future with little discernible shape. One difficulty was in reconciling participants' time constraints and potential technophobia (or Netphobia) with the ever-changing, ever-challenging Internet environment to which I was trying to introduce them. Another was to satisfy the needs of on-site trainers and support staff for materials which could be adopted, adapted and cascaded to others beyond the dozen or so attending the workshops each time. The route from task specification to task completion (not that it will ever really be complete - that's not the way of the networks nor the way of training!) was a challenge to ideas about teaching (training) and learning.

Initially there was a period of stock-taking, using paper and on-line materials developed during SOSIG's formative years. An on-line welcome page was loaded in a browser and bookmarked before participants entered the room and greeted them when they arrived. They were invited to browse through it to gain some WWW and Netscape background, something which seemed to challenge their ideas of what a workshop should be: they expected to be welcomed, introduced, led gently into the topic with a talk or perhaps a demonstration, not allowed to explore on their own. Some had dabbled before, thought they knew a lot and were expecting something a lot more sophisticated. Others felt lost: how could they get on and experiment when no-one

had taught them what to do? Many simply sat and stared at the text on screen and then, as time went by, at the screen saver. Others browsed through the comforting pieces of paper they had been given, hoping for guidance from there, but still reluctant to put fingers to keyboard or mouse. The know-it-alls clicked off into a bravado show of hypertext highjumps which did little to reassure their colleagues. A few of the newcomers caught on and read, understood and followed links, started exploration.

The aim had been to avoid giving them SOSIG "on a plate" but to provide a menu of ingredients with which they could experiment to find the most appropriate mix to prepare themselves for future learning, future exploration, with some guidance, some structure. In the main the challenge presented by this slightly unconventional learning experience failed them. Instead of prompting reflection, questioning, experimentation, it engendered resentment amongst the knowing, misunderstandings and misconceptions amongst the beginners: some thought the welcome page was SOSIG, some were unaware of what they were using - a browser- to view this information: was it a word processor? a text reader? How had it appeared on their screens? Worse still, their ability to come to grips with paper exercises later on and their confidence to proceed further were seriously affected. The unconventional start which should have set a positive note of enquiry for the rest of the workshop instead proved a barrier to learning.

The pattern of the workshops and the materials evolved gradually for a time as I tried to improve them with small adjustments, but eventually changed dramatically to take on a more conventional look which could still incorporate prompts for reflection, areas of challenge. A very traditional start of welcome speech, presentation with slides about the Internet and demonstration of how to load the browser and the on-line tutorial (a several page and quite complex development on the one-page welcome page) now leads participants gently into the recipe, but the emphasis is placed early on on self-paced learning, exploration, challenge within a structured framework. The on-line 'slides' provide more flexibility than their Powerpoint predecessors:- the levels of experience of each audience varies enormously, from one university to another, department to department and within the same workshop. The hypertext links in the slides provide for a longer, detailed 'talk' for newcomers, with the ability to bypass the basics and/or offer off-the-cuff demonstrations for a more experienced set. The presentation can be expanded in almost any direction dictated by the experience and interests of the audience. Within the traditional framework of talk and presentations and step-by-step exercises implying apprenticeship, acceptance of information learnt from the expert, lie semi-Socratic interventions which stop short of destroying all former knowledge in order to clear the way for future learning: at all stages participants are questioned and challenged, either to think further about what they are doing, to seek information and provide an 'answer' and prompted to

develop their own questions, their own enquiry.

Once the expected presentation is over, participants load a browser, enter the URL for the on-line tutorial and start exploring. Some are hesitant and follow slavishly what is on screen, but for most, the flexibility of the tutorial structure encourages them to set their own agenda, their own pattern for learning. After these initial explorations via the tutorial and a talk about and demonstration of SOSIG, participants are finally given the comforting pieces of paper many of them still crave. There is generally a collective sigh of relief at this stage: they have had a taste of freedom but they do not yet feel ready for total liberation. The pilot step-by-step exercises have been developed into a workbook and are interspersed with questions demanding reflection and further questioning in turn and to try to interrupt slavish adherence to instructions without understanding. Quizzes are provided to reinforce and test the newcomers' learning and to engage the more experienced. In both cases they are designed to illustrate how SOSIG and the resources it points to can be used to support teaching and research, to prompt lateral thinking, provoke consideration of searching and browsing strategies. Emphasis throughout is on consideration and development of the latter, on the different tools and techniques available, the different thinking that may be required depending on the design and content of the resource. This is followed through in further exercises involving other UK national services and international WWW search engines such as Alta Vista, Excite etc. Participants are encouraged throughout to use their own search terms or subject areas for browsing rather than sticking rigidly to the examples in the exercises. After the initial, traditional 'presentation' introducing the workshop and enough background to get them going, the rest of the day - and these are full-day sessions - is given over largely to participants' explorations. There is no requirement to use all of the workbook or to follow the sections in any particular order. Most are delighted to be able to set their own pace, to follow up their own lines of enquiry within the framework the workbook provides. During the second two thirds of the workshop my role is to respond to questions, interrupt occasionally with comments and demonstrations on issues which arise and offer collective or individual guidance if asked.

Requirements and the materials to meet them are constantly changing. Materials are frequently updated with screen shots and instructions, URLs and comment. I have to address participants' increasing levels of Net experience arising from extensions to campus networks and the increasing availability of graphical browsers on academics' office machines. Few sites now want coverage of using telnet to access WWW resources via Lynx, more want an introduction to HTML authoring. Web-based evaluation forms and questions and comments during workshops provide useful feedback in tailoring workshops and materials. Information is also collected at the workshops on participants' previous usage of the Internet and World Wide Web as part of the evaluation of

the effectiveness and usefulness of the Gateway and subject-based services in general, as well as of the training. Follow-up questionnaires and, in some cases, telephone interviews, seek to provide comparative data for usage after the workshops, to be analysed by a consultant employed under a related project. The most common comment on the workshop forms has been the usefulness of 'protected time' to explore, of not being forced through at a particular pace and of being allowed to follow up own lines of enquiry. Early on in the workshop participants are exposed to ways of recording and saving information found on the Web, through bookmarks, saving and copying and, in some cases, using electronic mail. Fewer now reach for the pen to write down URLs of useful sources they have found. Many copy bookmarks to disks to take away as a new starting point. If participants have not started following links or searching for information themselves by the middle of the second session I feel that the workshop has failed. The most successful sessions - from my own and from participants' points of view passed on through evaluation forms - are those where questions come thick and fast, where the text and graphics appearing on screens as I roam around the room are those I have never seen before, where participants call colleagues' attention to resources they have found, sometimes scampering excitedly around the room like children. Their enthusiasm for exploration, the discovery that amongst the abundance of networked information sources there are some which could prove really useful, that there are ways of finding and handling information which are not trivial or a waste of time, is a great joy. Even more so when they begin to think out loud, follow lines of thought on how they might incorporate some of the resources in their teaching, how they might introduce students to them. From then on the SOSIG has jumped off the plate, each participant leaves with a handful (or mindful) of ingredients which they can fashion into their own individual recipes for locating, using and perhaps in the future building networked information resources.

1. This paper has been presented at the CSS96/IASSIST conference at Minneapolis, University of Minnesota, May 12 - 19, 1996.

Nicky Ferguson and Lesly Huxley, Centre for Computing in the Social Sciences, University of Bristol 8, Woodland Rd BRISTOL BS8 1TN

E-MAIL:nicky.ferguson@bristol.ac.uk

CCSS home page at: <http://sosig.ac.uk/ccss/>

SOSIG at: <http://sosig.ac.uk/>

Data Services in a Non-Academic Environment

by Ineke A.L. Stoop¹,
Social and Cultural Planning Office,
The Netherland

The Social and Cultural Planning Office (SCP) is a Dutch government research agency. It contributes to policy making and appraisal in The Netherlands by producing a coherent picture of the state of social and cultural affairs. The production of such an overview requires an ever increasing amount of empirical data. During the 23 years of SCP's existence several hundred data files have been edited, stored and made available to the researchers. Data come from a large number of providers, such as Statistics Netherlands, commercial public opinion and market research firms, ministries and the Dutch data archive. Secondary use of data is strongly advocated Data services in a small non-academic organization differ by scale alone from data services in large academic institutions. Some problems, however, are universal.

The Netherland's Social and Cultural Planning Office

In 1973 the Dutch government founded the Social and Cultural Planning Office (SCP). SCP is an inter-ministerial scientific institute which conducts independent research on social and cultural aspects of government policy. Research findings are mainly intended for the Government, the Upper and Lower Houses of Parliament, senior officials of ministries, and for professional and academic staff in the public sector and academic circles.

SCP publishes the *Social and Cultural Report* every two years in September. An English version is available. The Report describes the social situation of the Dutch population and their attitudes to government policies on social and cultural matters. It covers health care, the social services, employment, social security, justice and criminal procedures, housing, education, leisure, the media and cultural affairs. Other publications deal with one of these subjects or with issues like the distribution of income and wealth, the economics of the public sector, the position of target groups or local policy matters.

SCP employs about 50 researchers, each an expert in one or several subject areas. In addition everyone has to be aware of current policy issues and to be versed in empirical research and data analysis. The department of Information Services supplies them with data, software, hardware and statistical consults.

Information Services at SCP

The department information services at SCP (Information and Automation, I&A) comprises a relatively extensive

group of people (10) for a relatively small number of researchers (50). Information services covers a wide area, though: information technology, communication, statistical consulting and data services.

Information technology at SCP stands for providing researchers and administrative staff with PC's software and information systems. All PC's are linked in a Novell network. Most users work in a Windows95 environment. Most statistical analyses are processed on dedicated PCs for running under OS/2. I&A personnel take turns in staffing a help desk and are also responsible for internal and external electronic communication (E-mail) and Internet facilities.

Statistical consulting encompasses a wide variety of activities, such as assisting researchers using SPSS, creating scales, and developing models, suggesting statistical methods, arranging courses and giving in-house demonstrations. When mounting surveys, methodological issues have to be taken care of, related to the determination of adequate sample sizes, weighting data, inputting missing data and wordy questions. The area of survey methodology might also be considered as part of the data services provided by I&A.

Data Services at SCP

Data services are part of the information services at SCP. Researchers need PC's to operate, statistical programs to manipulate, hypotheses and above all ... data to contemplate. As Sherlock Holmes said, in *The Adventure of the Copper Beeches*:

"Data! Data! Data! I can't make bricks without clay."

The first task of data services at SCP is determining what kind of data is required in a particular study. Researchers are urged to use existing data. For this purpose SCP has a wide range of data files available on-line:

- longitudinal series of population surveys from Statistics Netherlands on subjects like well-being and health, employment, education and expenditures;
- longitudinal series of surveys (partially) funded by SCP on facility use, time budgets, cultural changes and opinions;
- international surveys (Eurobarometer waves, International Social Survey Programme (ISSP) modules);
- institutional data (hospitals schools);

- one-time data (troubled youth, opinions on social security, environmental behavior);
- demographic data.

Data from Statistics Netherlands are obtained by way of the Scientific Statistical Agency, an intermediary institution acting as a go-between for (academic) researchers and Statistics Netherlands. Data from Statistics Netherlands are secured from disclosure by leaving out regional variables, minority group indicators and other identifying characteristics. In some cases only outdated survey waves are to be had, from other surveys recent data files are available. All things considered, Statistics Netherlands takes a co-operative line in providing research institutions with data. Other providers can be much harder to cope with. Sometimes data are withheld for strategic reasons or privacy considerations. Other times they are very expensive.

Mounting surveys funded by SCP start out; by selecting the research institute. In some cases a particular research institute has carried out a survey for years and is thus an obvious choice.

The preparation of a new survey requires a lot of time. Every detail requires attention. With respect to existing surveys, however, where comparability is paramount, routine set in as sampling procedures and questionnaires content are changes as little as possible.

Once data files have been obtained, they have to be edited and standardized. Weighting and imputation of missing data may have to be performed. Data files are all described in the Meta Information System SCP (MISS). MISS contains information on study, survey and variable level. Inclusion of the questionnaire and keywords may be possible in the future.

When data files are 'ready' they are all on-line available for researchers to analyze. In most cases these analyses are run on dedicated PC's using SPSS-X. SPSS for Windows, however, is gaining popularity. I&A comes into action again only for advanced data usage, e.g. when data files have to be matched synthetically or are molded into models. Of course, researchers often require assistance, either when they have questions about the survey or statistical or methodological questions.

During the years SCP has acquired a mountain of data. Sometimes data files are small, other times large with respect to number of cases, number of variables or complexity (multi-level data). Data files contain information on health, well-being, time budget, employment, social security, education, housing, media, expenditures, and so on. Data objects range from individuals and families by way of organizations and communities to postcode regions and EC-countries. Variables indicate opinions, describe behaviors or contain measurable quantities like expenditures, time

budgets, possessions (persons) or personnel, funding, housing (organizations). And finally, they are supplied by a large group of providers, as outlined above.

Quality Control

Sound government policy requires high standard research and high standard research requires first-class data. Unfortunately, social research in the Netherlands suffers from some major drawbacks, summed up by SCP as *the hole in the data layer*.

The first flaw in data collection is the poor response to surveys in the Netherlands, even in face to face research. Mind that there is no census in The Netherlands and that participation in official surveys is never mandatory. A response rate of 40% is no exception. Response to surveys is selective in a way not always to be remedied by weighting. The elderly, people from minority groups and inhabitants of inner cities are often underrepresented. Weighting by using postulation characteristics may partly compensate for this kind of nonresponse. Nonresponse may be related to the topic of the survey. In this case weighting will not always be possible. And finally, whereas SCP publications allegedly cover the entire population, actually, a lot of unsuitable cases are excluded from general social surveys. Inmates of prisons, hospitals, homes for the elderly and mental institutions are not interviewed. People who cannot see, hear, speak, write or read are frequently not able to participate and thus registered as nonresponse. Nonnatives may be under represented because of language problems. People either too young or too old are excluded as a standard procedure. Yuppies are never at home, the homeless cannot be reached, the mentally retarded may not understand every question and drug addicts live in another world. Though some of these problems may be solved by interviewing by proxy, a lot of people quietly drop out of the sample frame.

Nonresponse and noncoverage are tackled in three ways:

- firstly, take pains to boost response,
- secondly, find out in what manner respondents differ from nonrespondents,
- finally, compare results with data from other sources and, possibly, use data from public records.

Besides having an acceptable response rate, data ideally should come up to the following standards:

- have national coverage, and thus allow for regional comparisons;
- cover the entire Dutch population, and thus allow for comparison between groups;
- come from recurrent surveys (containing identical questions), and thus allow for comparisons over time;
- be fully documented.

Strengths and weaknesses

SCP publications are based on a firm empirical tradition. They nearly always contain empirical analyses. This gives SCP a clear advantage as compared to other institutions which mainly rely on material published elsewhere. Another reason why SCP publications tend to get a wide press coverage is that they generally bear upon current policy issues.

The main strength on SCP thus may be summarized as a tradition of empirical, policy relevant research. A major weakness originates from the size of the institute (rather small) and the range of subjects SCP covers (fairly wide). Researchers sometimes extend deadlines in order to create a better product. In other cases they desperately try to finish on time, putting up with data files not being analyzed thoroughly and only presenting cross tabulations, where as publications may be delayed, or data files may be analyzed insufficiently. Finally, analyses are not always documented in full. Reproducing analyses of former waves of a survey may thus require a laborious process of reconstruction.

Data services at SCP have their own strengths and weaknesses. After almost 25 years SCP has obtained a hoard of data on almost any social and cultural subject. Old files can be used to outline developments. New files can be used to illustrate current dilemmas. As said earlier, secondary use of data is advocated. Central storage of a single copy and on-line availability guarantee data integrity. Editing and standard format prevent errors. The main strength of data services at SCP, however, is that data services, information technology and statistical consulting are thoroughly integrated within one department. Conflicts of interest can be solved at the root, data collection can profit from methodological expertise and computational facilities are geared to a research environment.

Still all is not well. Firstly, much of the information about data files only exists in the head of the data people. Documentation is often incomplete. Questionnaires get lost or exist as a computer program only. And secondly, the tradition of secondary analysis does not come naturally to researchers, especially when they come straight from university. Most of them need time to get rid of the academic attitude: *get new data to test new hypotheses*.

Threats and opportunities

In its more complacent moods SCP might consider itself a victim of its own success. Competitors seem to be attracted to SCP research methods. Universities start proclaiming the use of secondary data and building large data bases to be used and reused by students instead of collecting data time and time again. Statistics Netherlands wants to get rid of a large series of unattractive publications filled with huge tedious tables containing detailed statistical data. Instead it wants to publish semi-glossy compilations of statistical data,

including pictures and sexy graphics. These new publications may appeal to the general public. SCP, on the other hand, was and still is extremely fond of huge tales and extremely detailed statistical data to be stored in SCP data bases and use for further analysis.

Another threat to SCP and its data services is the aforementioned hole in the data layer. Data deficiency is, among others, caused by a variety of restrictions of data use. Embargoes, privacy regulations and disclosure avoidance procedures limit the availability of data. The progressive use of public records by Statistics Netherlands, a development to be applauded in research itself, also threatens SCP data services, as administrative data will not be available for research outside the statistical office. As commercial data providers or associations of nonprofit organizations start to appreciate the strategic importance of information, the availability of data will fall off and the price will go up. And, as said earlier, the growing nonresponse rate in the Netherlands is a serious threat to data quality.

The final threat mentioned here will be considered an opportunity by a lot of people. In the old days, statistical programs, like SPSS required a set up which was generally stored for later use. Data files had to be stored centrally as local PC disks were much too small. SPSS for Windows does not require a set up; operations are performed directly on the data file. Analyses results can be cut and pasted into the final text. Recorded files for individual use can analyzing data, much more discipline is required to control quality. The development of explicit quality control procedures and tools at SCP is meant to help researchers to maintain this discipline.

Not all projects are bleak, though they require renewed efforts to keep data services at SCP up to par. Luckily, some new phenomena may help. The electronic super highway may be seen as an overload of use information. Still, when used prudently, it may offer useful tools for communication by electronic mail and web sites, for data collection and for data acquisition. Though SCP works for the Dutch market, international data can put this market into perspective. The Internet may help to obtain international data. And finally, the Internet may help in developing meta data at SCP, as it opens us international efforts and international standards for meta data and documentation. In its small way, SCP is starting to Weave the Web.

1. This paper has been presented at the CSS96/IASSIST conference at Minneapolis, University of Minnesota, May 12 - 19, 1996.

Social and Cultural Planning Office P.O. Box 37, 2280 AA Rijswijk, The Netherlands. phone: 31 70 319 8700; fax: 31 70 396 3000.E-mail: stoop@scp.nl; url: <http://www.scp.nl>



INTERNATIONAL ASSOCIATION FOR
SOCIAL SCIENCE INFORMATION
SERVICE AND TECHNOLOGY



ASSOCIATION INTERNATIONALE
POUR LES SERVICES ET
TECHNIQUES D'INFORMATION EN
SCIENCES SOCIALES

Membership form

The **International Association for Social Science Information Services and Technology** (IASSIST) is an international association of individuals who are engaged in the acquisition, processing, maintenance, and distribution of machine readable text and/or numeric social science data. The membership includes information system specialists, data base librarians or administrators, archivists, researchers, programmers, and managers. Their range of interests encompasses hard copy as well as machine readable data.

Paid-up members enjoy voting rights and receive the IASSIST QUARTERLY. They also benefit from re-

duced fees for attendance at regional and international conferences sponsored by IASSIST.

Membership fees are:

Regular Membership. \$40.00 per calendar year.

Student Membership: \$20.00 per calendar year.

Institutional subscriptions to the quarterly are available, but do not confer voting rights or other membership benefits.

Institutional Subscription:

\$70.00 per calendar year (includes one volume of the Quarterly)

I would like to become a member of IASSIST. Please see my choice below:

- \$40 Regular Membership
- \$20 Student Membership
- \$70 Institutional Membership

My primary Interests are:

- Archive Services/Administration
- Data Processing
- Data Management
- Research Applications
- Other (specify) _____

Please make checks payable to IASSIST and Mail to :
Mr. Marty Pawlocki
Treasurer, IASSIST
% 303 GSLIS Building,
Social Science Data
Archives, University of
California, 405 Hilgard
Avenue, Los Angeles, CA
90024-1484

Name / title

Institutional Affiliation

Mailing Address

City

Country / zip/ postal code / phone