

**FOR COLD:  
REPLY BY CSU SOCIAL SCIENCE RESEARCH AND INSTRUCTIONAL COUNCIL  
TO EAR COMMITTEE'S JANUARY 2002 REVIEW OF THE SOCIAL SCIENCES  
DATABASE ARCHIVE, CSULA**

**Elliott R. Barkan, CSUSB, Chair, SSRIC, 2001-2002  
[ebarkan@csusb.edu](mailto:ebarkan@csusb.edu) ; 909-880-5525 (o); 909-624-7679 (h)**

**February 12, 2002**

The SSRIC met on February 8, 2002, for its regular winter quarter meeting, at which time it extensively reviewed the EAR report. The following points address what we perceive as the central concerns and reservations expressed by members of EAR.

**Our general observation: Significant—and important--portions of the EAR report are based upon several false assumptions, which explains why EAR responded negatively to many issues related to the SSRIC/SSDBA proposal.**

**First**, the central issue here is not just the SSDBA's future role but insuring the continued funding for the principal data bases currently funded by the Chancellor's Office and managed by the SSDBA –ICPSR, Roper, and Field. It is understood by us that the objective now is to incorporate these databases into the PHAROS system of the CSU Digital Library and we are fully supportive of that move. ICPSR, for example, is the closest thing we have to a national social science data base. The SSRIC includes faculty representatives from each campus, and we feel strongly that the ICPSR, Roper, and Field Poll resources are invaluable resources that must remain available within the CSU system. Concerns about the SSDBA must be addressed separately, which we do below.

**Second**, there was a false assumption made that everything in the operation of the SSDBA, which EAR determined would be essential for its role in the CSU Digital System—along with the data bases--was already in place now, before it and the databases had even been recommended for inclusion. That was not possible and was never claimed to be the case—for example, no campus IP addresses (allowing for general access) could have been substituted for the passwords currently used by faculty until the whole conversion process had been agreed upon. Consequently, many criticisms expressed by the members of EAR were probably true but assumed something not yet in existence. We repeatedly acknowledged that there are alterations that need to be made in the present arrangement and we are confident that they can be done appropriately and in a timely manner. Indeed, important changes are already being implemented with the SSDBA, as are noted below.

**Third**, there is an assumption running through the EAR report that these raw social science data bases are somehow different from the kinds of supposedly readily accessible resources currently provided by the library, implicitly assuming that all such materials can be automatically used by students. Nothing could be farther from the truth. CSU libraries currently make available to students advanced articles in the hard sciences and theoretical articles, for example, in the social sciences, which cannot be used by students who have not had prior training with their instructors. There are many, many resources that students can access and download which they cannot now effectively utilize without prior instruction. In fact, portions of

the census are now available through the libraries but they cannot be readily used by students without prior training or assistance (because not all such raw data automatically comes in portable files). That is the function of the teachers, not the librarians.

We no more expect librarians to be able to explain how students should analyze or interpret an article in physics than we would expect them to explain how to do the SPSS analysis of the SSDBA databases. The criticism that social science data are somehow more complex or different than other materials made available by the libraries is unfair, for it draws a false dichotomy between these resources and others currently in use. Much of what the libraries provide students require that their instructors have first explained to them how to use those materials. There is no difference with the databases under review.

Let us be more specific: Consider an article in the *Journal of Mathematical Physics*. We do not expect the library staff to be able to read and understand the article but we do expect the staff to be able to locate the article using keywords that the user supplies and to be able to download or print the article—or to locate the article in a print journal. With the Census we do not expect librarians to be able to interpret data on the median price of housing in different cities, census tracts, or states. We do expect that librarians can assist students and faculty to locate the data in print or electronic form and to obtain a copy of the data.

Similarly with the social science data bases. Our expectation is that librarians will function in the same mode: helping users locate a data set and downloading the codebook, other instructions, and the data onto the user's machine or disk. At that point their role ends, for they are not expected to be able to run SPSS or another statistical package or to interpret multiple regressions. This is the role of the faculty or social science computer lab staff.

**Fourth**, another false inference is that the CSU is principally a system devoted to instruction (particularly undergraduate instruction) as if no research is done, but in fact a great deal of research is done by both instructors and students. That is how we train the students; that is how we attract top quality professors—by insuring them that they will be able to pursue their research agendas and thereby be able to produce works that will be incorporated into their instruction, provide support for their tenure applications, and validate their professional stature. We cannot endanger the funding of such vital databases as those under consideration here without seriously undermining the reputation of the CSU system. In fact, every single major university in the United States is a member of ICPSR. To suggest that the data bases are really just for graduate students and faculty also overlooks the key points that we do in fact have graduate students who need such materials; that we do have faculty who need such materials; and that such materials are frequently used by instructors throughout the CSU in their undergraduate classes as well as graduate ones.

**Fifth**, related to the above point is the mistaken assumption that a lot of materials provided by ICPSR, for example, are currently available for free. In fact, it is principally federally mandated and federally funded materials, such as the General Social Survey, the census and national elections surveys, that are available from other sites at no cost. Yet, even here, the Census Ferret program, for example, provides descriptive data, not data in an inferential format that would allow for empirical analysis (“manipulation,” as used in EAR's report). Thus, some

sources referred to only make available formatted text materials that do not provide the opportunity or possibility for in-depth research with them. And, even where data are made available for free, students and others often need assistance with the codebooks and organization of the data. In the final analysis, ninety percent of ICPSR's materials, as well as much of Roper's and all of Field's can only be obtained by subscription. Now, much of the libraries' target populations may be those seeking summary information (such as via "fact-finders"), but the presumption from the outset with regard to the databases under discussion has always been that users here would be those seeking raw data to manipulate, analyze, etc.--for instruction and research--and not only descriptive output (although even some of that can be gotten through links on the SSDBA site).

Related to this was the comment that some of our data sets, particularly the Field Polls, are not current. Such polls require adaptation for use on the campuses (in all systems, including UC) and we can only make available such polls and other data when they are sent to the SSDBA. These are not as current as, for example, newspapers, because they are fundamentally different types of materials. Processing raw data into SPSS portable files involves the time-consuming task of creating SPSS syntax files, which the SSDBA staff does. Nonetheless, every effort is made to keep them as current as possible, given the nature of the materials and our dependence on the various vendors.<sup>1</sup> Indeed, more recent Field Polls for early 2001 have been posted, and various other research studies are continually being added as they become available.

**Sixth**, EAR expressed concerns about costs and a fear that funding would soon dry up and devolve onto the libraries at the expense of other resources. But this is not an issue that SSDBA or SSRIC can solve; it is the responsibility of the Chancellor's Office, which initiated this process of incorporation, to provide the funds and insure that these will become a stable line item in the Digital System budget. We can only provide the support, not the moneys. However, we must emphasize that the CSU currently has substantial discounts in its database contracts because of our federation status; any plan that would decentralize or fragment the contracting process would prove very costly.

**Seventh**, While some individuals quoted in the EAR report complimented the SSDBA for its "general ease of use," some concerns related to the operation of the SSDBA were expressed by others. We must stress again that the SSDBA center developed precisely because there were instructional and research needs not being met in any other way. The concerns expressed by EAR focused on the need for training in order for students to be able to use these raw data sets, such as with SPSS or Excel, etc. But we now reiterate that the need for such training is, of course, recognized, but it is no different from all the other kinds of training students currently require in order to use other library materials, even bibliographic materials. SSDBA staff and members of the SSRIC have for years been providing such training for faculty,

---

<sup>1</sup> To illustrate the fact that we rely on the vendor, note the following response from the director of the Field Poll Institute, Mark DiCamillo, received 2/11/02: "Because the statewide primary elections are being held in March this year, we consider all data files conducted between September 2001 and March 2002 as pre-primary election data files, which are to be embargoed until after the primary election. This is similar to embargoes imposed on pre-election data files in previous election cycles."

who then have incorporated that training into their instruction so that students can use the data sets both in the classroom and in their independent work. This type training could continue, should continue, and should be expanded. As mentioned on page 9 of the EAR report, reference librarians could be given some training on how to search these data resources, just as they have been given assistance in learning how to access and use the other search engines added into the PHAROS system. The training could be adapted for student groups, as well.

Moreover, some components of the training materials will also be put on the SSDBA web site. In addition, even with the new ICPSR Direct—or with any free materials from ICPSR—no one should expect that the staff in Michigan will be available to provide direct assistance or training for students, faculty, and staff here in California. In fact, they have no intention of doing so. The SSDBA, therefore, remains an indispensable component as an adjunct, or supplement, to the reference librarians. And, given the size of the CSU system, we should not be relying on outside centers to provide such multi-level assistance; we must preserve, in effect, our own in-house provider of such services.

**Eighth**, some members of EAR expressed the concern that the SSDBA staff would not always be available and that students and others now electronically make use of library resources 24/7, but we are unaware of any other situation where library staff members are available 24/7. Reasonable availability of the SSDBA staff can be expected and, hopefully, at some point additional staff can be hired to expand the service hours. The SSDBA is also preparing improved documentation for using the data sets as students acquire more direct access to them and it will be developing FAQ files with models and representative outputs along with answers to basic questions for using these resources—similar in part to some of the materials available on the SSRIC's Teaching Resource Depository (TRD) web site. That site—accessible via the SSDBA home page--does contain, as well, some completed packages of data that can be directly utilized, a sort of middle road between the raw data sources and the descriptive information available from other providers (and from SSDBA).

**Ninth**, one member of EAR stated that “until it is easier to set up an account, I cannot support this...,” and another stressed that “IP address recognition is also a must for access.” The SSDBA has been limited to individuals with accounts and passwords, but all faculty, students, and staff (with faculty sponsors) on subscribing campuses have always been able to request such an account. And, the new automated account-creation feature makes this an extremely simple and rapid procedure. As suggested, the SSDBA has not currently been set up for direct access by students and faculty because that was not practical as long as individual campuses were subscribing (and a few did not) and almost only faculty were using the SSDBA. When the social science data sets are included in PHAROS and direct access by members of all campuses is required, the SSDBA will drop the password requirement. It will shift to IP authentication (domain name recognition) so that students, faculty, and staff on the campuses or accessing the libraries remotely (by first accessing the local campus library server as all now do) in order to use the various research engines will be able to do exactly the same thing with these data bases and the SSDBA site as they currently do with other electronic research tools.

In fact, it has already been pointed that the SSDBA is moving to a new search engine by April or May 2002, one that will include many more user-friendly components, such as bulletin

boards, on-line chat capabilities (including for specific courses), etc. It will also replace the current Iplanet pull down menu and provide a new front-end home page that will be much easier to use, with better help buttons to provide improved assistance for those trying to get started. Furthermore, the implementation of SDA (Survey Documentation and Analysis)—an online statistical analysis program—will likewise improve accessibility and “ease of use”. And, statistical and SPSS training documents have also been developed to assist individuals in downloading and working with SPSS files, descriptive statistics, tests of significance, and linear regression; these, too, will be made available to SSDBA users. Additionally, regarding codebooks, all hard copy codebooks are available from the SSDBA. Copies of the codebooks are provided to users at no cost. And, many of the hard copy codebooks are being converted into electronic format for more convenient access, particularly research studies obtained from ICPSR.

**Tenth**, when this new system is in place in PHAROS, additional staff members will eventually be needed by SSDBA to provide more extensive services than it now does because students will be accessing it and because all campuses will be participating (and contributing to the SSDBA revenues). It is believed that the enhanced system will also have a positive effect on the SSRIC members—and on the support they receive from their respective home institutions--and we are confident that they, too, will be more willing to help insure that the system, and the SSDBA in particular, runs efficiently and that the help and the training that people will need will be arranged.

**Eleventh**, and finally, since we have made it clear that we never suggested that everything was ready to be operationalized at once with SSDBA—knowing also that the process of COLD approval and Chancellor Office financing meant that full implementation was still some time away in any event—the SSRIC would like to endorse a proposal suggested on page 9 of the EAR report and expand upon it, namely by urging that

1. COLD reevaluate EAR’s report in light of the many comments made here,
2. COLD provisionally approve the inclusion of the social science data bases along with the support role of the SSDBA (the latter still funded by the individual campuses),
3. COLD then set up a subcommittee to work with the SSRIC and the SSDBA during the next academic year so that, with the assistance and input from expert librarians or other library staff, the SSDBA can be restructured or modified along lines that would be satisfactory to COLD and the needs of the PHAROS system,
4. COLD review in the spring of 2003 the recommendations and changes jointly proposed and implemented and then, if satisfied that its concerns have been met, give full approval for the formal inclusion of the social science data bases in the CSU Digital System along with the intermediary reference, assistance, training role of the SSDBA,

5. The Chancellor's Office provide assurances to COLD that the necessary funding for these databases will be forthcoming as a stable line item (bearing in mind that the individual campuses, and not the Chancellor's Office, separately fund the SSDBA specialty center and are expected to continue to do so), and that
6. The Chancellor's Office provide the SSDBA with adequate funds so that a qualified assistant manager can be hired to assist the present manager, Tony Hernandez, in adapting the SSDBA system to the needs of PHAROS upon consultation with the COLD subcommittee and along the lines of the improvements currently under way but very much in need of additional human resources.

This report was prepared by Elliott Barkan in consultation with members of the SSRIC. If COLD, or EAR, or Evan Reader, or Gerry Hanley have questions about the points raised here, or other ones, please do not hesitate to contact Elliott at his office—909-880-5525—or at home—909-624-7679—or by email: [ebarkan@csusb.edu](mailto:ebarkan@csusb.edu). Thank you,

Elliott R. Barkan, Chair, SSRIC  
Professor of History and Ethnic Studies, CSUSB  
President-elect, Immigration & Ethnic History Society  
Book Review Editor, *Journal of American Ethnic History*