

Access to Restricted Data in a Controlled Environment: The Michigan Center on the Demography of Aging Data Enclave

Michael A. Nolte
Senior Research Associate, The University of Michigan

Janet J. Keller
Research Associate, The University of Michigan

The Data Enclave of the Michigan Center on the Demography of Aging is a new facility designed to assist prospective users of restricted or sensitive data files -- files that cannot be distributed to the general public due to concerns about maintaining respondent confidentiality or because of licensing/use agreements which prohibit third-party redistribution. The Data Enclave provides a controlled, secure environment in which eligible researchers can conduct work that requires use of such restricted access data resources. This paper provides an overview of the organizational, legal, and technical issues associated with design and implementation of a secure facility. Disclosure limitation procedures used by Data Enclave staff when reviewing analysis results are also discussed.

To be presented at Session 237, "Enhancing Researcher Access to Confidential Data: Five Case Studies," Joint Statistical Meetings, Atlanta, Georgia, 8 August 2001. We acknowledge support from the Michigan Center on the Demography of Aging (NIA P30 AG12846) and the Health and Retirement Study (NIA U01 AG09740).

1 Introduction

1.1 Why Restrict Access to Data.

The past decade has seen unprecedented growth in the computing power and data storage capacity available to academic, government and commercial researchers. The new capabilities afforded by this resource growth have encouraged the creation of on-line collections of research data containing detailed information on multiple aspects of the lives of specific individuals. For example, Health and Retirement Study¹ collects information from each respondent on health status, cognitive condition, disabilities, retirement plans (including details on Social Security and pension wealth), family structure, employment status, job history, housing, income, and net worth. Such data collections are extraordinarily useful to researchers. If, however, all these data elements were to be made publicly available in an unrestricted fashion, the privacy rights of the study participants might be compromised. Even when direct identifiers (name, address) are removed, indirect match techniques that combine geographic information with unique demographic characteristics such as age, race, income, education, and detailed occupation/industry codes can be used by intruders attempting to breach respondent confidentiality.

The separation of respondent data into public and restricted categories can be an effective barrier against breaches of confidentiality. Public data files can be created in which potential demographic identifiers are either recoded to more general (less identifiable) levels or are masked entirely. Separate files containing more detailed demographic information can then be distributed under special, restricted conditions. In this paper we use the term "restricted data" to mean any file that cannot be distributed to the general public because of respondent confidentiality concerns or because third-party data use or licensing agreements prohibit redistribution.

The restricted distribution technique is an effective way to make available many types of respondent information that are of interest to researchers, but cannot be distributed to the general public. For example, the Health and Retirement Study obtains permission from participants for access to Social Security earnings and benefits records. The files derived from these administrative data records are highly sensitive and cannot be distributed to the general public. By treating these data as restricted, the study can grant access to researchers, and the privacy of respondents can still be protected.

1.2 Distribution Solutions

We will discuss three methods that allow the delivery of restricted data to researchers: licensing, remote access, and secure research facilities. Of these three, licensing is the most commonly used method, with secure research facilities coming rapidly into vogue. Access agreements that allow researchers to submit analysis requests via the Internet are still rare.

1.2.1 Licensing

The restricted file licensing process typically requires researchers to sign a contractual agreement with the holder of the restricted dataset defining the conditions under which research can be conducted and results published.² This agreement usually involves some type of financial commitment such as a security deposit (required by the Panel Study of Income Dynamics)³, or in the case of the Health and Retirement Study, a United States Federal grant that can be "held hostage."⁴

Although licensing agreements for restricted data access vary, the typical agreement contains the following elements:

¹ See the Health and Retirement Study Web site (<http://www.umich.edu/~hrswww>) for a full description of the study's holdings.

² See, for example, the National Center for Educational Statistics *Restricted-Use Data Procedures Manual*, at <http://nces.ed.gov/statprog/rudman/>.

³ The Panel Study of Income Dynamics (<http://www.isr.umich.edu/src/psid/>) allows researchers to use respondent geographic information and medical data through a special contractual arrangement.

⁴ See the Health and Retirement Study Restricted Data Web site (<http://hrs.isr.umich.edu:1041>) for details on the agreement under which restricted data are made available to researchers.

- a research plan describing how and why the restricted data are needed
- a data protection plan describing how the secure environment in which the data are to be used will be created and maintained
- a list of individuals who may have access to restricted data (research, technical, and support staff)
- institutional review board approval
- a formal contract between the researcher and the distributing agency that includes:
 - a description of the researcher's financial commitment
 - disclosure limitation rules
 - description of penalties for violations.

Restricted data agreements typically are time-limited and, in several instances⁵, enforced through on-site audits.

1.2.2 Remote Access

A second distribution method, which has been implemented at NCHS and elsewhere,⁶ is remote job submission. This technique allows users to submit analysis runs via the Internet to an applications server running at a remote site. The submission is reviewed at the remote site to ensure that it adheres to the conditions of use previously agreed upon by the researcher and the data holder. For example, an analysis request that produces a list of all respondent records will be rejected. The output produced by the analysis request is then reviewed, electronically or manually, in order to prevent disclosure of confidential information. If everything is in order, the results are returned to the researcher via electronic mail or some other communications method.

1.2.3 Secure Facility

In licensing restricted data, one problem that has arisen is that certain classes of user cannot comply with the terms and conditions of the restricted data agreement. For example, the Health and Retirement Study requires that researchers be current holders of a United States Federal Government grant. This in effect eliminates academic researchers without Federal funding, graduate students without a faculty sponsor, and potential users from the business world. Remote submission also has certain drawbacks. The process is subject to Internet-imposed security and capacity limitations, researchers may not have access to a full range of analytic tools, and implementation and maintenance of the server system requires skilled, dedicated support staff.

An alternative to licensing and remote submission is a third type of distribution method, the secure research facility. In a secure facility, visiting researchers are given access to a wide variety of restricted data under terms of use in which close supervision over user activities compensates for broadened access. The secure facility method is currently employed by the Health and Retirement Study and the Michigan Center on the Demography of Aging (MiCDA). Acting together, these research units of the University of Michigan have established a secure facility known as the MiCDA Data Enclave.⁷ The Enclave is designed to assist:

- Prospective users of HRS restricted data files, especially those derived from Social Security administrative data, who do not meet the requirements imposed by restricted data contractual agreements.
- Researchers who have special data analysis needs that cannot be met under the terms of a standard restricted data agreement.
- Junior faculty and graduate students from the University of Michigan and elsewhere.

⁵ NCES, HRS both use an on-site audit process to review researcher compliance with the researcher's data protection plan

⁶ Details on the National Center for Health Statistics Research Data Center are available at (<http://www.cdc.gov/nchs/r&d/rdc.htm>)

⁷ The MiCDA Data Enclave is funded by the Michigan Center on the Demography of Aging (NIA P30 AG12846), the Health and Retirement Study (NIA U01 AG09740), and the Michigan Retirement Research Center (SSA 10 P-98358-5). The Survey Research Center and the Population Studies Center of the Institute for Social Research also provide support for the operation of this facility. For more information, visit the MiCDA Data Enclave Web site (<http://micda.psc.isr.umich.edu/enclave/>).

1.3 The MiCDA Secure Facility Solution

The MiCDA Data Enclave secure research facility is located within the Institute for Social Research building on the campus of the University of Michigan. The Data Enclave is currently designed to serve two researchers at any given time. Data Enclave users are provided with office space and high-capacity personal computers that allow them to access the statistical analysis software, specialized application software, compilers and utilities that are necessary to manipulate and analyze restricted data files. The computers used by Data Enclave visitors communicate with a dedicated server on a network that has no physical connection to the Internet or to any other campus network. A unique password-protected profile is created for each user that gives him or her access to the restricted data files specified in the User Agreement.

1.3.1 Advantages

There are a number of advantages to the secure facility approach. First of all, the defined area of the facility allows the staff to exercise close control over the users: research results are scrutinized prior to export from the facility. At the same time, user access is improved, since the terms of access essentially depend on having a valid research interest and being able to meet the cost-recovery requirements.⁸ Finally, the controlled environment can be used for specialized research efforts that could not take place elsewhere, such as cross-category merges of restricted data files from various sources.

1.3.2 Disadvantages

The secure facility approach is not without its disadvantages. Startup costs are not trivial; setting up a secure facility entails a significant investment in time and money. Ongoing costs, both direct and indirect, are also high. Hardware and software must be maintained and upgraded. Staffing can be a problem, since employees with network systems, data management, statistical, and administrative skills must be recruited and retained. There is a certain amount of administrative effort entailed in negotiating with third-party data providers about how their data is to be used in the facility. Finally, the varied requirements of enclave visitors must be handled on a daily basis.

2 Establishing a Secure Facility: A "How-to" Outline

2.1 Define Project Scope

Establishing a secure data facility is fairly straightforward, provided that the necessary resources are devoted to establishing the ground rules for startup and subsequent operation. The first step is to define the scope of the project by conducting a needs assessment. This will allow the project planners to determine the target user community for the restricted data files to be made available through the facility. It will also provide the basis for estimating how many users can be expected to visit the facility in a given time period and what demands they are likely to place on the facility's software, hardware and data resources. Projected user demand will in turn help to define:

- Size and scope of data, software, hardware, and network resources.
- Staffing requirements.
- Space requirements.
- Time frame required for building a working facility.⁹

⁸ At the MiCDA Data Enclave, eligible users are defined as "faculty members of accredited institutions of higher education, students who are currently enrolled in an accredited graduate or undergraduate program and other researchers who wish to use MiCDA restricted data files but are unable to fulfill the contractual conditions for off-site access."

⁹ Establishment of the MiCDA Data Enclave required approximately two years: one year to plan and implement the facility and a second year to work out procedural issues. During this interval, a number of users from the Health and Retirement Study served as "beta testers" for the facility.

2.2 User Requirements

A list of user requirements will grow out of the initial needs assessment. The following table contains an outline of areas to consider in determining the data, software, hardware and network resources that the secure facility will provide to its users.

<i>System</i>	<i>Components</i>	<i>MiCDA Example</i>
<i>Applications Software</i>	Statistical packages	SAS, Stata, SPSS
	Office suite	Word, Excel, Access, PowerPoint
	Encryption	PGP, Microsoft Encrypting File System
	Utilities	McAfee VirusScan anti-virus software, TextPad text editor, Perl
	Compilers	Fortran, C++
<i>Data Products</i>	Inputs for user analysis	HRS/AHEAD Public and Restricted files
<i>Network</i>	Network/Server OS	Novell Netware V4.11
	File server	Compaq ProLiant 1600
	Printer	HP LaserJet 2100
	Hub and Cabling	HP 100/10 8 port
	Backup and archiving equipment	DLT Tape backup unit; CD writer; Zip drive
<i>Workstation</i>	Operating system	Windows NT 4.0, Windows 2000
	Special needs software	Vision aids, speech recognition, readers
	Hardware as required to support the components listed above	Gateway Pentium II 550 or greater

2.3 Staffing

The number and characteristics of secure facility personnel to be hired obviously depends on the mix of skills possessed by employees currently on your staff. Nonetheless, you will need to build a team that together has the following set of qualifications:

- *Network Systems Manager*: install and maintain server hardware/software; install and maintain workstation hardware/software maintenance; perform system backups; maintain network and workstation security.
- *Database Administrator*: install and maintain public and restricted data and metadata; archive user data; review and install user files; create user result export files.
- *Statistician*: conduct pre-export disclosure review.
- *Help Desk*: assist visiting researchers in using enclave facilities.
- *Administrator*: process user requests; provide general administrative support.

Note that depending on your local situation, additional personnel costs may be incurred if an oversight committee must be appointed in order to approve user requests. Also, you may wish to designate additional FTE fractions to deal with secure facility staff member vacation-time and sick-time needs.

2.4 Space Allocation

In academic environments there are few negotiations more problematic than those involved in obtaining a relatively large area that will be used on an intermittent basis to benefit a relatively small number of individuals. Nonetheless, it is important to realize that a secure data facility cannot operate in an area that is insufficient for its needs. At a minimum, a secure facility will require space for the following purposes:

- Secure location for network server, printer and hub.
- Office space for system manager, administrative support personnel, and user support people.
- Secure workspaces for visitors.
- Public workspace for visitors (i.e., for email, telephone calls, etc.).
- Secure storage space for facility records, archived data files and special types of data (e.g., videos, medical specimens).

In the case of the MiCDA Data Enclave, we have allocated two 8' x 12' offices for visitors and a 12' x 12' office which the Enclave manager shares with the network server, printer and hub. In addition we have provided cubicle space: one cube to be occupied by an Enclave support person and the other to be used as a public work area by visitors. One further point: when budgeting for startup costs related to office space, be sure to allow for:

- Renovations, especially HVAC and electrical modifications to accommodate computer equipment.
- Office furniture.
- Installation costs.
- Site changes to meet ADA requirements.

2.5 Operational and Startup Budget

Although startup and operational costs for a secure facility will vary widely depending on location, the experience of the MiCDA Data Enclave can serve as a guide. Startup costs for the facility amounted to \$30,000 in direct charges for equipment and site modifications. Approximately \$70,000 in indirect costs were incurred due to staff activities in developing the site plan, installing hardware and software, and establishing user procedures. Operating costs are approximately \$100,000 per year. This includes staffing, hardware and software maintenance, and space charges. It is important to note that the Data Enclave benefits from cost and resource sharing with the Health and Retirement Study, which on a yearly basis spends \$520,000 for the creation, updating, and distribution of restricted data products. It is possible that without this indirect support, particularly in the staffing area, operating costs for the Data Enclave might be considerably higher.

2.6 Operational Procedures

The final component of the startup process is to define the rules of operation that will apply to visitors at the secure facility. These operational procedures will also have an important bearing on how the duties of the facility's staff are defined. At a minimum, the operational procedures should define:

- The process by which user applications are received and approved.
- The legal rights and responsibilities of facility visitors.
- The installation and use of facility data resources.
- The rules for reviewing analysis results prior to their export from the secure facility.

The operational procedures established by the MiCDA Data Enclave, described below, are designed to meet these criteria.

2.6.1 The User Application Process

In order to maintain optimum access to a scarce resource, organizers of a secure data facility should implement a clearly defined user application process. This will assist the user in preparing for his/her visit, while providing the facility staff with information needed for planning. The main goal of the user application should be "no surprises."

In the case of the MiCDA Data Enclave, each prospective user is required to submit a research proposal containing the following information:

1. Cover letter.
2. Project Title.
3. Abstract.
4. Sponsoring Institution Approval: Users must provide full personal identification and institutional/organizational affiliation. If the prospective user is a student, a letter from the department chair or advisor stating that the applicant is a student working under the direction of the department must be provided.
5. IRB approval from sponsoring institution. If a local institutional review board is not available to the prospective researcher, then the University of Michigan Institutional Review Board can be used.
6. Current resume or Curriculum Vitae.
7. Dates of proposed tenure at the Data Enclave.
8. Funding source(s) for user project and for Data Enclave cost recovery.¹⁰
9. A detailed summary of the proposed research including a description of why publicly available data are insufficient to meet research needs, and/or why existing contractual methods for restricted data access are not feasible.
10. A complete list of the MiCDA data files being requested.
11. A description of user supplied data, if any, to be merged with the MiCDA data. This includes documentation, file layout, number of records, and restriction on MiCDA using the data (user-supplied data will not be released to anyone other than the prospective researcher) .
12. A list of software requirements.
13. A description of expected analysis results including (1) a list of research results to be exported from the Data Enclave and (2) storage format for exported data.
14. A list of special needs requirements (if any). The MiCDA Data Enclave will make all appropriate efforts to accommodate users with special needs.
15. Two signed copies of *Confidentiality Agreement Restricting Disclosure and Use of Data from the Michigan Center on the Demography of Aging Data Enclave*.
16. Two signed copies of the *Institute for Social Research Pledge to Safeguard Respondent Privacy*.

Once the research proposal package described above is received, it is reviewed to make certain that all items, including cost recovery details, are present and complete. The proposal is then reviewed for final approval or disapproval using the following criteria:

1. Scientific and technical feasibility of the project, including availability of data files being requested.
2. Approval by third-party data providers, if such approval is required.
3. Availability of resources at the Data Enclave: workstation time slots; file pre-processing schedule; pre-export disclosure review schedule.
4. Risk of disclosure of restricted information based on the user's description of expected analysis results. In particular, close scrutiny will be given to the list of analysis results to be exported.
5. Whether the goals of the proposed project are in accordance with the purpose of the MiCDA Data Enclave.

2.6.2 Legal Issues

It is important that researchers who apply for access to the resources of a secure facility recognize their duty to uphold the individual respondent's right to privacy by acting appropriately to avoid breaches of confidentiality. This can be done by informing applicants of what activities may or may not be performed during the course of their research work at the facility. This document should also include a short description of appropriate laws and/or regulations that can be applied in the event of user misconduct.

¹⁰Researchers using the MiCDA Data Enclave are charged for space and equipment rental and staff time necessary for supervision, disclosure limitation review, maintenance of computer facilities (including both hardware and software) and the creation and maintenance of data files required by the researcher. Daily charges, by user category, are: Academic/Government, \$200/day; Student, \$50/day; Other, \$500/day.

At the MiCDA Data Enclave, each user signs two documents as part of the application process: the *Confidentiality Agreement Restricting Disclosure and Use of Data from the Michigan Center on the Demography of Aging Data Enclave* and the *Institute for Social Research Pledge to Safeguard Respondent Privacy*. These documents define the rights and responsibilities of visiting researchers.

The *Confidentiality Agreement* clearly defines the rules with which Enclave visitors must comply:

1. Not to make copies of any files or portions of files to which the visitor is granted access.
2. To return to MiCDA Data Enclave staff upon request all materials (restricted and otherwise) with which the visitor may be provided during the conduct of the visitor's research.
3. To make no attempt to identify any household, family, person, establishment, or sampling unit.
4. To hold in strictest confidence the identification of any establishment or individual that may be inadvertently revealed in any documents, discussion, or analysis. Such inadvertent identification revealed in the visitor's analysis will be immediately brought to the attention of MiCDA Data Enclave staff.
5. Not to remove any printouts, electronic files, documents, or media until they have been scanned for disclosure risk by authorized MiCDA Data Enclave staff.
6. Not to remove from the MiCDA Data Enclave any written notes pertaining to the identification of any establishment, individual, or geographic area that may be revealed in the conduct of the visitor's research at the MiCDA Data Enclave.
7. To allow inspection of any material the visitor may bring to or remove from the MiCDA Data Enclave. Data Enclave staff may prohibit the removal of any material, including written notes, from the MiCDA Data Enclave.
8. The visitor also agrees to hold harmless and indemnify MiCDA and the University of Michigan, its agents and employees, for any claims of breaches of confidentiality arising out of the visitor's research and to pay MiCDA \$10,000 for each violation (defined as failure to abide by any section of this agreement or any accidental or intentional violation of privacy of any contributor to any MiCDA data resource) of this agreement.

The *Institute for Social Research Pledge* serves the additional purpose of notifying the visitor that adverse professional consequences may result from misuse of Enclave resources. By signing this document, the visitor certifies that he or she will abide by all ISR policies on safeguarding respondent privacy as a condition of continuing collaboration and association with the ISR, even if the visitor is not an employee of the ISR (e.g. a student, visiting scholar, or outside investigator).

2.6.3 Data Resource Management

Although a secure data facility must be organized to provide tight control over access to restricted data resources, its primary goal is to provide researchers with access to sensitive data so that they can perform analysis and publish their results. But what if the output of an analysis run or the interaction of certain restricted data files provide clues to respondent identities? The solution to this problem is to implement a set of data resource management procedures that apply to all stages of data use within the secure facility. The overall goals of these procedures are to:

- Prevent disclosure of confidential information.
- Reduce the likelihood of respondent re-identification.
- Provide useful data resources to researchers.
- Ensure that the results of the review process are acceptable to data users and providers.

There is no universal set of procedures that defines how these goals can be met. At the federal level, agencies have evolved a variety of solutions to meet their own unique needs. For example, *Working Paper 22*¹¹ describes disclosure limitation methodologies for twelve Federal agencies; there are eleven separate solutions to the problem. One resource that we found useful in developing data resource management procedures was the *Checklist on Disclosure Potential of Proposed Data Releases*, developed in draft by the Interagency Confidentiality and

¹¹ *Report on Statistical Disclosure Limitation Methodology (Working Paper 22)*, Federal Committee on Statistical Methodology, Office of Management and Budget, May 1994, p. 40

Access Group.¹² This document includes three sub-lists -- microdata, demographic tabular data, establishment tabular data -- that contain questions designed to help reviewing authorities determine the rules under which information derived from confidential data can be released to the public. Informed by these documents, MiCDA Data Enclave staff members have developed a set of data management procedures that, in our opinion, are workable and meet the goals described above.

2.6.3.1 Data Inputs

The data management process begins with inputs -- the files that are available to Enclave users. All data products available for use by Enclave visitors are subjected to a variety of modifications in order to ensure respondent confidentiality.

- MiCDA Data Enclave microdata products are directly or indirectly based on a survey sample rather than a population census. In the case of HRS microdata files, each respondent represents approximately 2000 individuals in the target population. This severely limits the opportunity of an intruder to match a given case to a population list.
- Public file variables containing industry, occupation, and geographic information have been bracketed and/or masked. Details are available only as restricted data.
- Microdata files derived from SSA administrative data (Wage/Self-Employment Income, Earnings and Benefits Data) have undergone rounding and top-coding in accordance with the governing Memoranda of Understanding.
- Direct respondent identifiers such as name, street address, Social Security number, Medicare identifier, Medicaid identifier, etc. have been removed from all public and restricted microdata products.
- Restrictions are placed on access to geographic detail information by users of SSA administrative data.
- Data elements derived from sample design components such as PSU, segment, and line information, are not made available to users.

2.6.3.2 Access Limitation

Although the purpose of the Enclave is to provide researchers with access to the information that has been removed from the public files, in order to further reduce disclosure risk Enclave visitors are given access only to the restricted data files specified in their User Agreement. Access limitation is enforced by both network software and Enclave staff surveillance.

2.6.3.3 Pre-Export Disclosure Review

The final step in protecting restricted data files is for all analysis outputs to undergo a disclosure limitation review conducted by Enclave staff members. Details of the pre-export review process follow:

- Data Enclave users may export only statistical summary information (frequency tabulations, magnitude tabulations, means, variances, regression coefficients, and correlation coefficients) that does not permit the direct or indirect identification of any individual person, family, household, employer or benefit provider.
- Export of microdata files or analysis output containing information at the respondent or household level is not allowed unless specific permission has been obtained from the Michigan Center on the Demography of Aging and all relevant restricted data providers.
- Users may not remove any analysis output that can potentially identify respondents, sampling information, or geographic areas below the level of Census Region, either directly or inferentially.
- Tabulations may be exported, but are subject to the following rules:
 - Magnitude Data: Cell/stratum $N > 3$ is required.
 - Frequency Data: A threshold rule of marginals > 5 and cells > 3 will be applied to all tabulations.
- Cross-Category merges (e.g., State with Social Security Earning and Benefits Data) are allowed only under special circumstances.
- Analysis results containing merged area data based on geographic information may be exported if there is no direct identification of geographic areas, or if geographic areas are reported using the same grouping characteristics as public files. Any analysis result that directly or indirectly identifies geographic areas below the level of Census Region, as a row or column heading, may not be exported. Geographic

¹² Checklist on Disclosure Potential of Proposed Data Releases (Draft), Interagency Confidentiality and Data Access Group, Interest Group of the Federal Committee on Statistical Methodology, 1997

information may not be used in conjunction with files derived from Social Security Administration administrative data without written permission from the Social Security Administration.

- Disclosure review rules for special merges (i.e., the cross-category merges referenced above) are based on negotiations among MiCDA, restricted data provider (e.g., HRS), and the researcher, using the review principles stated above as a starting point.
- All analysis output is subject to disclosure review by Enclave staff members who, in consultation with restricted data providers, have ultimate authority over whether a given set of analysis results may be exported.

3 Some Final Thoughts

I'd like to conclude with a few observations on what we went through at the Michigan Center on the Demographics of Aging in setting up our own secure facility. Since the secure facility concept is a relatively new one, there was no "Dummies Guide to Data Enclaves" that we could use as a cookbook. Taking the route of least resistance, we did the easy things first. We negotiated for space (in retrospect, not so easy), renovated offices, bought software and hardware, and set up our private network. Two years later, this decision seems to have been a good one. By setting up our new private network to match the configuration of the existing Institute for Social Research public network, we were able to buy software and hardware "off the shelf." This parallel development decision also gave us access to the knowledge and skills of in-house networking experts. In the same vein, the expertise of Health and Retirement Study staff members in dealing with restricted data served as a springboard for defining the environment in which Enclave visitors would work.

Perhaps the most difficult task that we faced was how to deal with procedural issues. We needed to set the ground rules for researcher access to Enclave resources, with respect to inputs (third party data resources), the visitors themselves (application review procedures, legal issues) and outputs (pre-export disclosure review). Since each of these areas involved multiple constituencies, the development process for our organizational procedures was of necessity an iterative one, entailing considerable effort in negotiating a set of rules that would be acceptable to all interested parties. The outcome of these negotiations, described above, should provide a useful starting point for secure facility developers who wish to follow in our footsteps.

Given that the only constant in the computing world is growth, we are already planning for future changes and enhancements to the Enclave. If the demand for Enclave services exceeds our current projections, we will need to obtain more physical space for expansion. Another space issue that we will need to address if utilization levels increase is how to provide additional work area for support staff. With respect to security enhancements, we would like to begin encryption of network traffic and to provide automatic encryption of file folder contents. Since Novell does not supply encryption services, we are investigating the possibility of using Windows 2000 client and server software features such as IPsec and the Encrypting File System. We also plan to investigate biometric tools for regulating access to Enclave workspaces. As the various projects associated with MiCDA continue to produce new data products, we anticipate a corresponding increase in Enclave offerings. Anticipating this growth, we have tripled server capacity and are prepared for further expansion if necessary. Additional data products may also require new types of application software, a consideration that also factors into our plans for future expansion.

In conclusion, the secure facility approach has provided MiCDA with an effective, if somewhat costly, solution to the problem of distributing restricted data files. The physical setup of workspace, hardware and software was straightforward; defining procedures for use was more difficult, but well worth the extra effort. The MiCDA Data Enclave is now in full operation, so if you're interested in using the facility for your research, if you'd like to employ a similar solution for your own restricted data needs, or if you are just curious about our operation, please visit our Web site: <http://micda.psc.isr.umich.edu/enclave>.