FINAL REPORT NSF-LAMP PROJECT

Identifying Where Technology Logging and Monitoring for Increased Security End and Violations of Personal Privacy and Student Records Begin

American Association of Collegiate Registrars and Admissions Officers

American Association of Collegiate
Registrars and Admissions Officers
One Dupont Circle, NW
Suite 520
Washington, DC 20036
tel: (202) 293-9161
fax: (202) 872-8857
www.aacrao.org

Published by:
The following material is provided by the American Association of Collegiate Registrars and Admissions Officers (AACRAO) to its membership as a source of general information. This material represents AACRAO's general understanding of compliance requirements under current law as of the date of publication noted in this material. This publication is not intended to serve as a substitute for legal advice, and we recommend that in developing any specific campus policies, institutions seek the advice of counsel.

FINAL REPORT NSF-LAMP PROJECT

Identifying Where Technology Logging and Monitoring for Increased Security End and Violations of Personal Privacy and Student Records Begin

Prepared by:

VIRGINIA E. REZMIERSKI

NATHANIEL ST.CLAIR, II

A Report to the Digital Government Program of the National Science Foundation

This research was made possible by funding from the Digital Government Program of the National Science Foundation. Although the topic of computer logging and monitoring in colleges and universities may have seemed somewhat peripheral to the focus in this program at the time the application was received, the project proposal raised many issues regarding data collection and surveillance technologies that are also important to digital government. The program review committee recognized the relevance of the Logging, Monitoring, and Privacy (LAMP) Project research and supported its funding. The research team acknowledges with gratitude their foresight and interest in this specific set of questions and issues.

Contents

Acknowledgements

Project Staff

In addition to the authors of this report, others have contributed to its success: Two part-time research assistants, Jamie Hine and Aiyana Thompson, committed time and talent to the interview and data collection process. Their communication and analytic skills and their knowledge of the critical issues were extremely valuable. They helped to design the study, develop and pilot the interview instrument, and complete many interviews. We gratefully acknowledge their contribution to this project.

Virginia E. Rezmierski, Ph.D.
Project Director
Former Director,
Office of Policy Development and Education
Adjunct Associate Professor, Gerald Ford School of Public Policy and the School of Information
University of Michigan

Nathaniel St. Clair, II, B.S.E.
Research Associate, College of Engineering
University of Michigan
Currently: JD student, Marquette University Law School

Jamie Hine, B.S.
Research Associate
University of Michigan
Currently: JD candidate, University of Michigan Law School and MPP student, Gerald R. Ford School of Public Policy, University of Michigan

Aiyana Thompson, B.A.
Research Associate
Currently: MPP student, Gerald R. Ford School of Public Policy, University of Michigan

Project Advisory Board

This research was guided during its development and implementation stages by a group of technical specialists and policy experts: the project advisory board. Its members brought important content and perspective to this project, providing individual expertise and guidance at relevant stages, including technical, editorial, and policy expertise as needed. They contributed viewpoints from the large national constituency of organizations they represent. Their commitment to service and to better understanding of information technology and policy issues helped focus and guide this study.

Mark Luker, Ph.D.
Vice President
EDUCAUSE

Barmak Nassirian
Associate Executive Director
AACRAO

Richard Rainsberger, Ph.D.
Director, Admissions and Records
Central Virginia Community College

Eugene Schultz, Ph.D.
Adjunct Professor
University of California

Barbara Simons, Ph.D.
Former President, Association of Computing Machinery

Aline Soules
Scholarly Communication Librarian
Business School
University of Michigan

Dana Walker
Assistant Director, Associate Services
AACRAO (Advisory Board member 2000-2001)
Berkeley Laboratory

Pilot Study Participants

The research team acknowledges the valuable input provided by staff members at the University of Michigan where the pilot study was completed. These individuals were patient in helping us understand terminology and the technical aspects of computer logging. They helped isolate concepts and refine the project design. We express appreciation for the time and expertise of:

Mark W. Giuffrida, Associate Director
Computer Aided Engineering Network

Paul Howell, Systems Research Programmer
Office of Executive Vice President for Finance

Dave Stempien, Technologist
Information Technology Division

Kurt Hillig II, Technologist
Information Technology Division

Gloria Love, Systems Consultant II
College of Literature, Science and Arts

Karen Pachla, Systems Administrator III
College of Literature, Science and Arts

K. Scott Vowels, Systems Administrator III, M-Care

Lynn Zhang, Systems Administrator III
College of Literature, Science and Arts

Technical Focus Group

Technical assistance also was received from a specialized focus group of three individuals who provided clarification on several questions that arose following the critical data collection process. Sincere gratitude is expressed to these individuals for their valuable time and effort throughout this endeavor.

Clifford Collins
Network Security
Oarnet

Eugene Spafford
Director, CERIAS
Department. of Computer Science
Purdue University

Paul Howell
Senior Systems Research Programmer
Office of Executive VP for Finance
MAIS, Security Office
University of Michigan

Editorial Review Board

Finally, a collection of technical and policy experts agreed to serve as the editorial review board for this research project. The purpose of this group was to respond to the early draft of this final report and to comment on the soundness of the authors' conclusions and recommendations. This input proved extremely valuable. Recognition is rightfully due and sincerely given to the following group for refining this project's final report.

Clair Goldsmith, Vice President Information Technology
University of Alabama-Birmingham

Marjorie Hodges Shaw, Director Emeritus, Cornell Computer Policy & Law Program; Scandling Scholar, Warner Graduate School of Higher Education & Human Development University of Rochester

Rob Kling, Ph.D., Professor
School of Library and Information Science
Indiana University

Kathy Kimball, Director of Security
Pennsylvania State University

Christine Pruess, Project Manager
Information Technology Services
University of Iowa

Joseph M. Saul, J.D., President
Communications Technology Consultancy

Robert Ellis Smith, Publisher
Privacy Journal

Executive Summary

The Digital Government Program of the National Science Foundation funded the Logging and Monitoring Privacy Project, LAMP. The project examined the types and extent of computer logging on a sample of college and university campuses. Researchers investigated the purpose of such logging, how much and what kind of information was being collected, how that information was being used, and how many people had access to it. In describing the types of information being collected, the researchers explored whether such information could provide the identities of specific individuals and how many steps it would take system administrators to construct the identity of an individual student from the log data. Finally, the LAMP team sought to determine whether the kinds of log data being collected constitute "education records" as defined by the Family Educational Rights and Privacy Act (FERPA) and if various uses of the data violate the privacy rights of students under the act. The overall objective of the study was to provide information for college and university personnel regarding logging activities and to inform them about when such activities might require substantial limitations, management guidelines, and additional training requirements or policies.

Sixteen geographically clustered colleges and universities participated in this study. The institutions varied in size, experience in the use of electronic networks, and mission (public or private, two-year or four-year institutions). The first group of participants was selected system administrators from the participating schools. Fifty-seven system administrators participated in the study-an average of approximately three from each school. On each campus, system administrators who were considered the "most knowledgeable" concerning computer logging were identified; they "knew the most about computer logging on networks, administrative systems, central systems, and within a single large college." A member of the project team interviewed each system administrator individually. A standard questionnaire was used to collect data. System administrators provided data regarding their major systems, most in the Unix family of operating systems, and the vast majority of which were providing applications services to the institutions as a whole.

The second group of study participants included registrars. Fourteen registrars from the sixteen institutions participated. Registrars were selected because they are responsible for implementing FERPA on campus and possess the expertise to interpret the FERPA regulations. In responding to a preselected set of six scenarios describing various logging events and resultant data, they provided information regarding privacy and the protections afforded to "education records." Also included in this second group of participants were two experts from the Department of Education's Family Policy Compliance Office. Like the registrars, the experts were asked to respond to the scenarios. Matches between the registrars and the experts were analyzed, as were all of the responses to the five questions that followed each scenario.

This report discusses the implications of the different responses relative to the law and to each of the issues investigated, *e.g.*, whether log data constitute a record under FERPA, whether sharing constitutes a violation, whether the situations described in the scenarios qualify as legitimate educational interest under the law, whether data collection in itself violates the law, and whether the situations described are appropriate educational uses of data.

The report highlights specific and obvious findings of the research and conclusions relative to the participants and their training and experience. The authors discuss factors that seem to act as barriers to participants receiving sufficient training in security, data protection, and FERPA.

In this report, the authors discuss specific and obvious findings relative to campus practices. Data confirm that logging is being done at a high rate and that there is a desire to do more. Data also reveal that primarily default and enabled logging are occurring and that they are being done primarily for the purposes of security, systems and network maintenance, and operations management. The authors describe the types of data being collected in logs and the ease with which system administrators can use those data to identify specific individuals, usually without additional authorizations. They provide data regarding incidents in which individuals are the targets of logging activities and data are collected for potential investigation or tracking of activities.

This report details findings relative to education records, specifically the strong agreement among registrars and the experts that the log data described in the scenarios do qualify

as education records under FERPA. The report discusses differences in responses and the lack of consensus relative to questions pertaining to the sharing of data and the definition of legitimate educational interest.

Beyond specific conclusions, this report also addresses broader issues raised by the research. For each of the issues, recommendations are offered. Issues include the dynamic pressures being placed on system administrators and registrars; the increasingly networked electronic environments of campuses are making it increasingly difficult to protect systems and individual rights. The report identifies a condition of responsibility without adequate information for system administrators, one of decision responsibility without full technical understanding for registrars, and the overload and frustration of professional technical staff in general. The authors believe these factors contribute to confusion and perhaps even liabilities concerning the relationship among logs, education records, and student rights under FERPA. The report discusses inadequate protection from unwitting acts and suggests that inadequate protection in fact results from the absence of policies and the lack of guidance regarding the handling of logs and log data.

Three specific levels of logging are recommended; these levels vary according to the amount of individual information collected, the number of people given access permissions, and other characteristics. The levels are offered as potential guidelines for colleges and universities as they address logging activities on their campuses. The report discusses investigation and pursuit without sufficient collaborative consultation, the dangers of taking a low road to fair information practice, and the creep of surveillance technologies as major issues requiring immediate discussion and action on campus today.

Finally, the report provides several appendices, including a list of the participating schools, the system administrator data collection instrument, and the registrar scenarios and data collection instrument. It is hoped that this report will result in increased protections for education records and additional research regarding these important issues.

Introduction

This report provides information about a project funded by the National Science Foundation and awarded to researchers at the University of Michigan. The project was entitled "Identifying Where Logging and Monitoring for Increased Security End and Where Violations of Personal Privacy and Student Records Begin." The acronym, LAMP (Logging And Monitoring Privacy) will be used throughout this report to refer to the project. The report is organized into six sections:

- Section 1 provides the introduction, background, and purpose of the study.
- Section 2 describes the project methodology.
- Section 3 provides the results of data analyses from system administrator participants.
- Section 4 provides the results of data analyses regarding logging and monitoring processes.
- Section 5 provides the results of data analyses regarding student records from registrars.
- Section 6 offers recommendations and conclusions.

Background

Use of information technology on college and university campuses has grown exponentially during the past ten years. Computers are an integral part of the educational process for nearly every department within an education institution. They are required for student work in most professional preparation programs. Electronic communication, considered by many to be a critical component of campus life, is the most widely used aspect of technology in higher education. Networked services that facilitate collaboration with colleagues on different campuses and in different states and countries are growing in importance. The value of networked courses, offered at a distance from the campus, is being evaluated. Information resources made available through the World Wide Web have increased faculty, staff, and student access to publications, collections, and a wide variety of other valuable knowledge. The administrative business processes of colleges and universities also have become dependent on networks and information technology.

As more and more of the mission and business of institutions of higher education are carried out over networked information infrastructures, the importance of providing a secure environment for individual and corporate data, communications, and research and teaching information grows. Three important aspects of security must be realized:, confidentiality, validity, and integrity. Confidentiality becomes an important aspect of systems operations because sensitive or proprietary information must be protected. Validity of data within systems must be ensured because modified or false data can drastically affect the lives and reputations of individuals as well as the business functions of institutions. Systems that are dependent on network infrastructure must have integrity-they must be protected from unauthorized intrusion or tampering that can result in damage to resources, denial of services to the campus community, or civil suits against the institution.

These systems are vulnerable to abuse and misuse because of their complexity. Often, adequate security protections are not put into place when systems are installed. Complex systems frequently are not documented in ways that will ensure stability and appropriate and consistent use over time, as systems age and computing personnel change roles. Networks, not adequately protected, provide potential portals into institutional systems for those who seek to steal or abuse the resources that reside there.

College and university environments also are special by nature, increasing the complexity of their information systems and the inherent vulnerability of networks and systems. As open systems-systems that require the free flow of information and data from multiple sources-they must be configured to send and receive large quantities of information. The commonly used security devices that are designed to control the sending and receiving of data through firewalls and highly regulated authentication and authorization schemes have less value in academic environments, where such regulations would restrict the free exchange and rapid sharing of information. College and university communities are vulnerable to unwitting as well as purposeful abuses of network and information systems. After all, it is the role of students to experiment, explore, learn, and create.

Technology is changing so rapidly that the overall nature of the networked environments being created within institutions and corporations is changing too. Within the past few years, colleges and universities have moved from dependency on mainframes, where data protection and management

tasks were controlled more centrally, to distributed computing systems. , where data and information flow among servers, mainframes, and desktop machines over electronic networks. Different operating systems are tied together with varying degrees of compatibility. Data protection and management tasks in these environments depend on controls and responsible processes at different levels within data systems, at different authority levels within the organization, and often at different physical locations on campus.

The job of system administration in these complex settings is extensive and demanding. Ensuring the compatibility of systems alone is a major responsibility. Upgrading systems to maximize usability by employees is another significant and seemingly unending task for system administrators. System administrators are responsible for identifying and fixing vulnerabilities within operating systems and applications when new systems are introduced and/or changed. This is happening at a rapid pace. System administrators who were considered expert on one hardware platform often are expected to manage three to five additional platforms on which they may have received little or no training.

System administrators often are required to manage these increasingly complex environments without the benefit of sufficient time to fix systems, document processes, plan for changes, support users, or take advantage of in-service training opportunities. Information technology personnel also change positions often, a result in part of the availability of new employment opportunities, of stress and frustration on the job, and of the frequent reorganizations that plague information technology organizations.

System administrators have begun to increase the amount of machine and network logging and monitoring[1] they do on their systems. They have done this primarily to manage the systems, to understand the flow and loads of traffic on systems, to be aware of potential problems in systems functioning, and to be aware of potential abuses of services. They use various utilities built into systems to collect data, usually for later analysis. The collected data are stored in data files called logs. Such logs contain information about individual machines, network connections, the time and date of connections, and the numbers of individual attempts within a time period to access various services. They also may contain information about the processing load on each machine, the number of individuals "signed on to" or using particular services at one time, and patterns of use. High volumes of traffic over computer networks, increased network complexity, and increasing numbers of abuses of resources have had the effect of increasing system administrators' desire to observe and control the traffic and the user behaviors on the networks and systems they administer.

When systems data are collected in logs, such data include information that itself or when matched with other data can be used to identify individuals and their behavior patterns. As college and university environments increase the number of functions that are networked, the ability to create an increasingly complex picture of individual activities grows. What may begin as logging activity to protect the efficient and effective functioning of one system can become targeted data collection and surveillance of a specific individual. If that individual is abusing a system or committing an abuse against another member of the community, such data may prove highly valuable for investigative or disciplinary purposes. Yet it also may be an inappropriate invasion of privacy or an attempt to injure another person. Very little legislation and no guidelines exist to regulate the extent to which a student's actions may be monitored on a private or public university network and what notification and due process must be assured.

Nevertheless, one law is highly relevant: the Family Educational Rights and Privacy Act of 1974 (hereafter referred to as the Act or FERPA).[2] (See Bibliography for additional resources.) This law was written to afford students and their parents (in the case of minor students) certain rights to the protection of their education records. The American Association of Collegiate Registrars and Admissions Officers (AACRAO) has provides a summary of the rights and responsibilities of colleges and universities relative to this Act: [3]

"**The primary rights afforded are:**

- the right to inspect and review the education records;
- the right to seek to have the records amended; and
- the right to have some control over the disclosure of information from the records.

Education institutions and agencies are required to conform to fair information practices. This means that persons who are subjects of data systems (*i.e.*, students at an institution) must:

- be informed of the existence of such systems,
- have identified for them what data about them are on record,

1 Logging is the systematic collection of data regarding the computer system and/or network and about individuals using or abusing the resources.

2 Federal Register: 6 July 2000 (Vol. 65, Number 130) 33 CFR Part 99 Family Educational Rights and Privacy; Page 41851-41863, Final Rule.

3 The AACRAO 2001 FERPA Guide, 2000, Washington, DC,p.1.

- be given assurances that such data are used only for intended purposes,
- be given the opportunity to request an amendment or correction to their records,
- be certain that those responsible for data systems take reasonable precautions to prevent misuse of the data, and
- know that the institution will respond reasonably when an alleged misuse of, or access to, data is brought to the attention of those responsible for data systems."

In providing certain rights to students, the law designates education record as those records directly related to a student and maintained by the institution. If a record is directly related to a student—*i.e.*, identifiably associated with a specific individual—and if it is retained by the institution in any form (*e.g.*, handwriting, print, tapes, film, microfilm, microfiche, any form of electronic data storage), it is an education record under the law, and the student is afforded certain rights. Exceptions are stated in the law, *e.g.*, instructional or supervisory records in sole possession of the creator, law enforcement records created by a law enforcement unit for the purpose of law enforcement, employment records made and maintained in the normal course of business and relating exclusively to an individual's capacity as an employee, medical records not disclosed to anyone other than those providing treatment, and institutional records relating to a person after his departure from campus and after he is no longer a student at the institution.

The law states that institutions must give students annual notification of their rights according to the Act. They must notify students of what the institution has categorized as public or directory information as well as of their right to protect that information from disclosure. They must be informed of who has access to their education record information and for what purposes. While many specific rights and responsibilities are described within the Act, it is not concerned exclusively with records that contain grades and/or course information. Rather, the Act defines any data that are specifically identifiable to an individual student and that are retained by the institution as an education record (with certain exceptions). Many questions therefore must be asked regarding the relation of FERPA to the log data that system administrators increasingly feel compelled to collect. These questions include the following:

- Exactly what logging are system administrators doing on the systems they manage?
- How much logging is needed to secure and manage the systems?
- What is the purpose(s) of such logging?
- What actions are being taken to expand or to make logging more comprehensive?
- How many steps/actions must be taken to get from logs to an individual's identification?
- Does the person to whom the data refer know about the existence of the log?
- What protections are in place for the sensitive information collected in logs?
- Do data contained in system logs create a threat to the privacy of students on campus?
- Do log data constitute an 'education record' according to FERPA and thus require protections provided under the law?
- Is the collection of such information a violation of FERPA?
- Is the sharing/disclosure of such information a violation of FERPA?
- Would access to such log information qualify under 'legitimate educational interest' within the framework of FERPA?
- What training and support are needed for system administrators in this area as they fulfill their increasingly complex responsibilities?

Purpose

The purpose of the Logging and Monitoring Privacy (LAMP) Project was to:

- Examine and define stages of technical logging and monitoring necessitated by the need for more aggressive security within technology environments;
- Define conditions under which collection, storage, or disclosure of personal information from computer logs constitutes a potential violation of student record privacy law; and
- Define the interface between needed security logging and monitoring and potential privacy violations.

Results of this project will be incorporated into an educational white paper for distribution to colleges and universities. It is hoped that this project will stimulate additional research in this important area of information technology policy.

Methodology

Pilot Study

The LAMP project team recognized that systems logging and monitoring activities were relatively new on campuses. As with so many areas associated with the fast emerging technologies, the team expected differences in the use of terminology, in the knowledge of how different technologies function, and in the requirements for data protection under the law. To better understand these differences, pilot interviews were held on the University of Michigan campus. Senior staff experienced in technology and data management were asked to discuss issues, processes, and logging procedures. Through these discussions, the team explored current practices and future capabilities for logging and monitoring. The team refined the data collection instrument and clarified procedures for identifying the most appropriate participants for the study. The team also gained a better understanding of the complexities that exist in bringing together technical, legal, and policy experts for meaningful communication.

Selection of Participating Colleges/Universities

The budget for the LAMP project was modest. The research team knew that the identification of appropriate individual participants and the investigation and description of logging and monitoring activities at each participating college would require careful and direct communication. The research team therefore identified personal, on site investigation and data collection as important to the project design. To leverage travel funds, the team gathered data from colleges and universities that were easily accessible. They also gathered data from institutions that could be accessed during the research team's other required regional and national travel.

The selection of participating schools thus was not a randomized process. Sixteen schools were invited to participate, and all agreed to do so. (See Appendix A for a list of participating schools.) In response to encouragement from the project's advisory board, the team selected clusters of schools on the West Coast, on the East Coast, in the South, and in the Midwest. Several Michigan schools were added to the Midwest cluster because of their proximity.

The LAMP project team selected colleges and universities of different sizes (enrollments ranging from 1,300 to 43,000), offering instruction at different levels (two-year degree to graduate degree granting), and representing the designated regions of the country. The team also selected schools experienced in implementing information technologies as well as those newer to the process. With the exception of the midwestern cluster, which included four additional Michigan schools (as noted above), three colleges/universities within each region were invited to participate.

Data Collection

The data gathering process focused primarily on two groups of individual participants. Data regarding logging and monitoring activities on systems and networks were collected from system administrators. Data regarding definitions of student educational records, appropriate use of information, and the implementation of FERPA were collected from college and university registrars.

Selection of Study Subjects: System Administrators

The first group of individuals selected to participate in this study included network and system administrators at the participating colleges and universities. The study was described to Chief Information Officers and/or IT Security Officers. The team asked these administrators to identify three primary participants: the individuals on their campuses who "know the most about logging and monitoring on networks, on administrative data, and on central computing systems."

During the pilot phase, the LAMP team found that significant logging activity was occurring at the college and department level. Therefore, for each university, the name of the individual who knew the most about college-level logging in one of the largest colleges was also requested. Not all of the participating schools had individual colleges. Therefore, for some of the participating schools, only three representatives provided data.[4] An average of 3.5 representatives provided data from each participating school. Twenty-eight (28) percent of the study participant sample were network administrators, 28 percent were administrative computing managers, and 28 percent were system administrators for centralized services. Sixteen (16) percent of study participants were managers of computing at a major college within the institution.

4 The organizational structure of one of the universities made it impossible to reasonably separate the network, administrative, and centralized services from one another. Consequently, only two participants at that institution were identified.

Selection of Study Subjects: Registrars

At each college or university, one person is identified as the "steward" for student data. The title given to that person is *registrar*. The registrar, in conjunction with employees in the registrar's office, is responsible for designing procedures and policies to manage student data in compliance with federal law—specifically, FERPA. To gain expert input on the study questions regarding education records and the appropriate use of student-related data, the project team identified the registrar as the expert authority within each participating institution. Registrars therefore constitute the second group of participants in this study. Of the 16 registrars invited to participate in the study, 14 agreed to do so.

Research Materials and Tools

The LAMP project data were gathered through in-person interviews with the system administrators and through telephone interviews with the registrars. Once personal contacts with each of the system administrator participants had been established, the in-person interviews were completed. A pre-designed questionnaire guided the interviews with system administrators and helped maintain consistency. The questionnaire included questions about system administrator training and experience, the types of network and systems being used, the primary functions of those systems, the "personally identifiable information" the logs yield, and the intended purpose of logging activities. (See Appendix B for a copy of the system administrator data collection instrument.) The system administrators were cooperative and informative, often describing in detail their roles, systems, the scope of operations, and their frustrations with the constraints of time and resources.

Phone interviews were conducted with the registrar participants to collect data regarding education records and appropriate data use. The project's advisory board urged the use of scenarios as a mechanism for data collection. Each participating registrar therefore was asked to review a set of six scenarios and questions sent prior to the interviews. The scenarios selected for the LAMP study were chosen by the project's advisory board and staff members from a set of nine. The scenarios provided a range of issues and potential data uses for the registrars to contemplate and judge. Each scenario described a real-life activity of data collection, access, or transport drawn from actual incidents but stylized to provide anonymity. Following review of each scenario, the registrars were asked to answer five questions. (Appendix F contains the six scenarios and the accompanying questions.)

Overall, the LAMP team collected and analyzed data from 16 participating colleges and universities, from 56 individual participants (administrators of computer networks, central, administrative, and college-level systems), and 14 registrars. It collected information about logging and monitoring for a total of 76 systems.

Data Analyses: Participants

Who Are the Participants

Having identified the schools and the appropriate participants from networking, central, administrative, and college-level systems administration, the project team sought to understand the nature of the experience and training of these institutional staff. (See Appendix C for sample system administrator job description.)

Degrees

The data revealed that the 56 responding system administrators had a wide range of educational backgrounds: Two reported that they had completed their formal training at the high school level; ten had associate degrees from two-year colleges; thirty-one had bachelor's degrees (two had earned a B.A, and twenty-nine had earned a B.S.); eleven had master's degrees (two had earned an M.A. and nine had earned an M.S.); and one had a Ph.D. A wide range of academic majors was represented, including music, math, history, accounting, political science, psychology, and the more "expected" majors of computer science and information systems management. Aggregate data showed that 67 percent of the participants had science-related degrees and 9 percent had degrees in the arts.

Certifications

Technical certification programs are offered to computer employees by vendor organizations and by regional and national professional organizations. Such programs provide, usually at a high cost, a series of intensive training opportunities and testing that culminate in a certificate that designates the individual as a skilled person on a particular hardware platform, *e.g.*, Novell, Unix, etc. The LAMP project team explored the extent to which study participants had obtained technical certifications. Eight of the 57 respondents had obtained certification.

There are many explanations for the small number of certificate holders among respondent system administrators. First, because the programs are reported to vary greatly in quality, they may not be of interest to most system administrators. Second, for many colleges and universities, the programs are too costly. System administrators report that the cost of the program is usually greater than the training resources allocated to them. They also report that time constraints prevent them from taking advantage of training opportunities. Finally, many certification programs focus exclusively on one hardware platform, while system administrators increasingly are expected to manage multiple operating systems and their interactions. Several respondents indicated that though they themselves have had limited opportunities to gain certification, they encourage their technical staff to do so if funding is available and when the training is appropriate and of good quality.

The project team had questions about some of the responses in this area and therefore sought input from the project's technical focus group.[5] With regard to technical certification, the team asked, "Why do you think that so few of our respondents have sought or obtained technical certifications?" The focus group indicated that employers typically do not reward employees who obtain certification independently. Although certifications vary in quality, many will increase their holders' marketability. Many universities do not encourage their computing staff to acquire certifications because the institutions would be unable to compensate them at market value. Many system administrators are encouraged to attend conferences and seminars in lieu of technical training. Given these circumstances, computing staff continue to receive meager training, and institutional officers are less likely to contend with requests for salary increases from staff who acquire technical certification.

Experience

The LAMP study participants had significant professional experience. Researchers asked, "How long have you been in the field?" and "How long have you been in your current job?" Possible responses were (1) less than one year, (2) 2–5 years, (3) 5–10 years, and (4) more than 10 years. Nearly 90 percent of study participants—51 of 57—had been in the field more than ten years. Nine percent (5 of 57 respondents) had been in the field of computing for between five and ten years. Forty-five of the 57 participants also provided their actual numbers of years in response to these two questions. Table 1 shows the participants' years in the computing field. Almost 50% had been in the field for between 14 and 22 years. Twenty-two percent had been in this field for between 25 and 42 years. Respondents had a mean of twenty years in the field of computing/information technology.

5 A team of three individuals who are regionally or nationally known for their technical expertise identified to provide technical focus and guidance to the LAMP project.

TABLE 1. NUMBER OF YEARS IN FIELDFIELD

Years in Field	Frequency	Percent	Valid Percent
2-5	1	1.8	1.8
5-10	5	8.8	8.8
10 years	51	89.5	89.5
Total	57	100	100

TABLE 2: ACTUAL NUMBER OF YEARS IN FIELD AND JOB BY RESPONSIBILITY CATEGORY

Responsibility Category	Actual Number of Years in Field	Actual Number of Years in Job
Network Administrator		
Mean	18.36	6.27
N	11	11
Std. Deviation	6.71	5.37
Minimum	11	1
Maximum	33	20
Central Systems		
Mean	17.77	7.54
N	13	13
Std. Deviation	6.18	8.39
Minimum	9	1
Maximum	30	31
Administrative Systems		
Mean	22.15	10.64
N	13	14
Std. Deviation	9.81	9.79
Minimum	7	1
Maximum	42	28
College-Level Systems		
Mean	21	11.25
N	8	8
Std. Deviation	9.52	9.44
Minimum	5	1
Maximum	35	24
Total		
Mean	19.76	8.83
N	45	46
Std. Deviation	8.06	8.42
Minimum	5	1
Maximum	42	31

Even assuming that new or expanded areas of responsibility and changing jobs entirely are common occurrences in the field of computing, project team members were surprised by the number of respondents who said they were unsure what their title meant. Respondents also said that their title and responsibilities have changed often due to departmental reorganizations. While 33 percent of respondents had been in their current job for more than ten years and 19 percent for between five and ten years, nearly 50 percent had been in their current job for less than five years (23 percent for between two and five years, and 25 percent for less than one year.) Data pertaining to the actual number of years in current job confirmed this high change rate. Thirty-nine (39) percent of respondents had been in their current jobs for between one and three years. (57 percent for less than six years).

Analysis of the actual number of years in the field and in current jobs by administrative category, *e.g.* networks, central, administrative, or college level (see Table 2), shows even stronger evidence of high levels of experience. The means for respondents who were knowledgeable about logging on networks were 18.4 years in the field and 6.3 years in their current job. (The medians were 15 and 6, respectively). The means for central system administrators were 17.8 years in the field and 7.5 years in their current job. (Medians were 16 and 3, respectively). The means for administrative systems respondents were 22.2 years in the field and 10.6 years in their current job. (The medians were 23 and 7, respectively). Respondents from college-level systems administration had means of 21.0 years in the field and 11.3 years in their current job. (The medians were 23 and 8, respectively). The number of years of experience and the respondents' knowledge of computing were impressive. Many had been involved in the use of information technology on campus since the introduction of mainframes for central data processing and had aided the move to networks and distributed computing.

Training for Data Protection

The LAMP project sought to identify the logging and monitoring that were occurring on college campuses and their relationship to privacy protections under FERPA. The team therefore sought to determine what kind of training system administrator respondents had in security for the operating systems they administered; in data protection for the data they handled; and in protection of student records under FERPA. The team asked the following questions:

- Have you taken courses in systems security for the systems you administer?

- Have you taken courses in fair information practice, data access, and data protection?
- Have you taken courses in compliance with FERPA for handling student records?

The results were surprising. Only 21 percent of respondents had taken courses in security for the systems they administer. Twelve (12) percent had taken courses in fair information practice, data access, and data protection. None had taken courses in FERPA and the handling of student records. These respondents have the responsibility and ability to access and alter data stored in their systems. They also have many years of experience and significant knowledge, both of which may provide them with information about these topics.

Given the responses the study team received, its members considered the possibility that the research question regarding training may have been poorly worded, yielding faulty data. Asking respondents if they had taken courses may have encouraged them to apply too strict and formal criteria to their training. In fact, few formal courses on these topics are available to system administrators. Professional associations such as USENIX, SAGE, CIS, and others offer high-quality training, but the cost of the training remains prohibitive for many. Some commercial vendors provide courses in specific aspects of security or data protection. But they, too, are usually costly, offered in conjunction with professional conferences and thus requiring trainees to incur registration, travel, and lodging expenses as well. Like the certification programs discussed previously, they often cost more than what is allocated for college/university system administrator training. It also is true that few courses are offered at all. Technology has changed so rapidly that few people possess the expertise in systems security and data and records protection necessary to teach such classes.

Because it was important to understand the extent of the respondents' training in these areas, the study team performed secondary data collection on this topic. Hypothesizing that over their careers, respondents may have received short duration training in security and data protection through seminars at professional conferences, the team asked participants to estimate the number of hours of training they had received in each of the three areas (security, data protection, and FERPA). Thirty-six of the 57 study participants responded to this request for secondary data. The secondary data were not unlike the original data. Table 3 provides descriptive statistics showing estimated number of hours of training by respondent category for each of the three subject areas. Respondents representing network administration had the largest number of hours of training in security: a mean of 58 hours. Those representing administrative computing had the largest number of estimated hours in fair information practice, data access, and data protection: a mean of 54 hours. No group had a mean of more than 2.5 hours in training on FERPA and student record protections; the median for all groups was zero. Data analysis showed that the number of hours of training in FERPA most

TABLE 3: ESTIMATED NUMBER OF HOURS OF TRAINING BY RESPONSIBILITY CATEGORY

Responsibility	Estimated Number of Hours of Instruction in Security	Estimated Number of Hours of Instruction in Data Protection	Estimated Number of Hours of Instruction in FERPA
Network Administrator			
Mean	58.1	10.7	0.3
N	11	11	11
Standard. Deviation	47	29.8	0.6
Minimum	0	0	0
Maximum	120	100	2
Central Systems			
Mean	28	7.5	0.8
N	10	10	10
Standard. Deviation	49	14	1.9
Minimum	0	0	0
Maximum	160	40	6
Administrative Systems			
Mean	28.1	54.2	2.5
N	10	10	10
Standard Deviation	36.3	60.9	6.2
Minimum	0	0	0
Maximum	100	160	20
College-Level Systems			
Mean	25.4	2	0
N	5	5	5
Standard Deviation	35.8	2.3	0
Minimum	0	0	0
Maximum	88	5	0
Total			
Mean	36.9	20.7	1
N	36	36	36
Standard Deviation	44.1	41.4	3.5
Minimum	0	0	0
Maximum	160	160	20

likely were gained from in-house consultants and/or on-campus administrative meetings.

Data analysis showed that these individuals, identified as knowing the most about logging and monitoring on networks, central, administrative, and college-level systems in the participating schools, represent impressive numbers of years in the field. They have a wide range of formal educational backgrounds (as represented by terminal degrees), a small number of specific certifications on the systems they administer, and a small number of estimated hours of training/instruction in security, data protection, and/or FERPA. These respondents, many of whom likely obtained much of their knowledge on the job during the time when operating systems were being developed and networks were being implemented, may be recognized as pioneers in computing on campus. Then and now, they have worked with time constraints and shortages in both human and financial resources. Courses in the areas of security, data management, and FERPA have not been—and still are not—readily available to them. Because of the power and responsibility to manage systems and data, many have done extensive reading and study on their own. Yet precisely because of respondents' power and responsibility, the study team concluded that the need for additional training in data protection and FERPA for these and other system administrators on college and university campuses is substantial.

Data Analyses: Logging Processes

What Logging and Monitoring Are Being Done?

A primary objective of the LAMP project was to explore, describe, and understand the amount and type of logging and monitoring taking place at colleges and universities. Once the team was satisfied that the instrument and the interview process were sufficiently customized, it began its series of in-person interviews with project participants. (Analysis of the personal information gained from system administrators was discussed above, in Section 3 of this report.) Next came the collection and analysis of systems and process information. Seventy-five (75) percent of participants provided data regarding one of the primary systems they administered; twenty-two (22) percent provided data on two systems; three (3) percent provided data on three systems.

Kinds of Systems and Primary Functions

Data analysis showed that of the systems discussed by the 57 project respondents in this study, the largest percentage of systems being used (55 percent) were Unix operating systems (Unix, HP-UX, AIX, Solaris, BSD). Next in frequency (18 percent) was the Windows family of operating systems (NT, Windows 95, 98, & 2000). Mainframes (AS400 and OS390) made up 15 percent of the sample described. Network systems, such as CISCO and Novell, accounted for 8 percent, and Linux for 4 percent.

The team asked respondents to identify the primary function of the machine about which they were reporting. The machines were functioning primarily as applications servers for login, mail, domain name, and file services. Another portion of the machines were providing mainframe and infrastructure monitoring functions, such as load and traffic analysis. The team also asked to whom the services were being provided: department, school, university, state, or other.

Approximately 71 percent of these primary systems were running enterprise services for the institution as a whole. Eleven (11) percent were providing services for individual departments, 18 percent for a particular college or school within the university, and 9 percent for the state as a whole. Many administrators (11 percent) reported that the services provided were for consortia of organizations—some within and some outside of the university. The last two categories, constituting 20 percent of responses, provide important information. Some of the members of the consortia being served by the universities and some of those services going to the states represent K-12 populations—children considered minors. The study team hypothesized that if log information from these machines qualifies as "education records" under the regulations of FERPA, then the protections afforded to such records, for minors in the K-12 populations as well as for students within the college or university, may become the responsibility of these system administrators.[6]

Sensitivity of Data and Needed Protections

Researchers asked respondents to describe the level of sensitivity of the information contained within their machines. Specifically, they asked:

- "On a scale of 1 to 4, how sensitive do you think the data on this system are?" A score of 1 indicated "not sensitive" and 4 indicated "extremely sensitive".
- "On a scale of 1 to 4, how much protection should such information have?" A score of 1 indicated "no protection," and a score of 4 indicated "extensive protection."

The mean sensitivity rating was 3.2, indicating a high estimate of data sensitivity. Approximately 70 percent of those reporting on applications servers scored their data as very or extremely sensitive (3 or 4). Approximately 60 percent of those referring to systems whose primary function was mainframe or infrastructure monitoring also scored the data as very or extremely sensitive.

It is doubtful that a consistent reference point existed for the system administrators as they responded to study questions about data sensitivity. Those respondents most accustomed to considering levels of data sensitivity—those whose job it is to manage administrative data on mainframes and large databases—appeared to base their judgment of data sensitivity on what they knew about characteristics inherent in the data itself, *e.g.*, importance or sensitivity of the data to the individual to whom it referred. Others seemed to base their judgment of data sensitivity on their opinion of what

6 For students not yet age 18, FERPA rights are granted to parents unless the student is attending an institution of higher education.

someone could do to the system if the data were accessed. The LAMP team concluded that given a probable difference in respondents' understanding and definition of data sensitivity, these data must be interpreted cautiously.

Do Administrators Log?

Analysis revealed that for 96.2 percent of the systems reported on in the LAMP study, at least one logging capability was turned on: Log data are being collected. Logging, analysis of logs, and monitoring of systems were considered important parts of system administrators' responsibilities. The project team found that system administrators were logging for three primary purposes: (1) security, (2) network and systems maintenance, and (3) operating system management.

1. Security was the most prevalent reason given and included watching for abuse of systems, tracking attempted accesses by individuals in order to identify unauthorized access attempts (also known as intrusion detection), monitoring authorization processes to ensure that only those individuals with adequate and appropriate permissions were allowed to access data that require special authorizations.

2. Network and systems maintenance included activities such as troubleshooting, traffic analysis, and watching for system errors.

3. Operating system management ranked third in terms of purpose of logging. This category included activities for debugging systems, watching for limitations of software, providing system back-ups, and keeping a history of configurations.

Forty-four (44) percent of the logging was reported to be for security reasons; 28 percent was for purposes of network maintenance; and 28 percent was for purposes of operations management. These data confirmed previous information from system administrators that logging is an important tool for providing efficient and stable operation of systems as well as the necessary security.

The LAMP data show that 96 percent of respondents perform logging on the machines they administer. Further analysis shows not only that logging is being done, but that more than half (59 percent) of respondents would like to do more logging. The research team asked the system administrators to identify what barriers2 exist that keep them from doing more data logging. (See Table 4.)

*Note that a high percentage of respondents identified "other" as a barrier against more logging. "Other" often referred to resources such as additional staff, file space, and adequate central systems for such activities.

TABLE 4: BARRIERS AGAINST MORE LOGGING

Barriers[7]	Percentage Identifying Barrier
Time	57.7
Equipment	50
Other	38.5*
Policy	11.5
Personal Ethics	9.6
Law	7.7
Knowledge	3.8
Skill	1.9
Authorization	0

Default, Enabled, or Scripted Logging

The LAMP project staff members asked each system administrator about the type of logging used: if logging was default; if they had to enable functionality to accomplish logging; and if they customized or scripted the logging functions to meet their specific needs. Researchers assumed that the use of those terms was sufficiently common that respondents would have little difficulty answering the questions. However, the team found that different administrators use the terms "default", "enabled", and "scripted" differently and that they also mean different things depending on the type of system being administered. The meaning of the responses to these three questions therefore is unclear—or, at best, suggestive. The technical focus group was asked how it would define the terms "default," "enabled," and "customized/scripted" and on which systems default logging is provided. (See Appendix D for their definitions.).

Respondents varied greatly in their knowledge of logging functions. While some respondents were extremely familiar with the different logging capabilities of different systems, others were uncertain as to whether logging was enabled on some of the systems they themselves managed. The matter of logging functions is further complicated by the fact that systems are changing rapidly, and vendors package new logging functions in the default configurations of their updated systems. Again it became apparent that system administra-

7 "Barriers" were meant to refer to various types of preventives to action. They might be perceived as positive, such as rules, policies, or guidelines that need to be consulted prior to action and that need to be adhered to, perhaps slowing response time. Or they might be perceived as negative, such as rules, policies, or guidelines that need to be consulted prior to action and that are seen as impeding a rapid response and resolution of problems. Respondents were given several examples, such as knowledge, time, skill, authorization, equipment, policy, law, personal ethics or other, from which to choose in responding.

tors have difficulty keeping adequately informed about the systems they manage.

Analysis revealed that approximately 72 percent of respondents used the default logging function on their systems. A high percentage—approximately 82 percent—either enabled or disabled certain functions on their systems to obtain the desired type of logging. Some system functions were turned off because logging those functions would have provided too much data to be useful or practical. For still other systems, logs were not considered useful.

Another study question asked about customization of logs. Approximately 64 percent of respondents customized their logging processes by writing instructions (scripts) for the level and type of logging desired. The scripts typically were designed to instruct the computer to organize the log output to facilitate searches for specific information or specific machines. Scripts also were used to automatically send mail to or otherwise notify one of several specific individuals when particular patterns of use or abuse were found in the log data. Analysis showed that 14 of the 42 respondents said they did so to facilitate searching as well as to notify appropriate personnel; eight did so to enable specific rules that help maintain the stability or security of machine functions.

Data Collected and Steps to Identification

Having established that 96 percent of the described systems run logging functions, researchers asked about the type of data such logs yield. They also asked, "What and how many steps[8] would you need to go through to get from these data to a person's name?" If data are collected in large quantities to manage the computer and network systems on which college and university communities increasingly depend, and if such data contain information that readily identifies individual members of the community, then the issue of its appropriate protection under the law is increasingly important.

Table 5 shows the frequencies of different types of data yielded by computer logs. Ninety-three (93) percent of the reported logs provided date and time stamps for the transactions logged. Approximately 80 percent of the reported logs provided the user ID of individual users. Sixty-eight (68) percent provided the IP address of particular computer systems being used. Fifty-seven (57) percent reported the Domain Name from which a transaction originated. Fifty-four (54) percent provided other non-specified data, such as mail address and account information. (Appendix E provides figures showing the data yielded in system logs.)

TABLE 5: PERCENTAGE OF LOGGING YIELDS BY TYPE

Data Yield	Percentage
Date & Time Stamp	93.2
User ID	79.7
IP Address	67.6
Domain Names	56.8
Other	54.1

The terms "date and time stamps," "user ID," "IP address," and "domain names" are defined in Appendix D;, they are defined here as well because of their importance in understanding the relationship of log data to education records under FERPA.

- Date and time stamps are computer-generated pieces of information that tell the system administrator specifically when the user accessed the computer network and for what length of time he or she stayed on the network. By combining this information with other information types listed below, the system administrator often can determine with great specificity a user's current physical location, actions, applications accessed, duration of activity, and any changes made to existing files and other resources.

- A *user ID* is a login alias usually specified by the individual. It identifies all of the computing transactions of the individual associated with it. While a user ID is not generally an individual's first or last name, it is directly associated with identification and correlates through the computer account at the institution with a specific individual. It is an alias for a person's name while he or she is on the network.

- An *IP address* is the address given to a specific machine when it is attached to the network. Given an IP address, a system administrator can attempt to trace a specific computer connection or transaction back to the machine that was used for the action. While that machine typically is used or owned by one individual (except in the case of machines located in public computing facilities), it is not possible to guarantee that the individual who typically uses or owns that machine performed the action seen at that IP address. It is possible that another person used the machine with or without the permission of the person who registered it at the time it first was connected to the network.

8 The term "steps" as used in this study is synonymous with "actions."

- A *domain name* is the name associated with the section of the network on which a particular machine resides. It is used by routing mechanisms for electronic mail delivery and the transport of other packets of information from one network to another. With a domain name, a system administrator can attempt to identify the network segment from which a particular transaction originated, the institution from which the transmission originated, and the management authority under which the network is being operated.

System administrators typically have the responsibility and the ability to access information from many university sources. Because of their responsibility to supervise systems and networks, most are given "root" access on the machines they administer. Root access is required for operations management. It allows individuals to change the machines in many ways: to access and change functions, to change data and files, and to search and combine data from many different sources. With their role comes responsibility not only for machine operations but also for data protection. Responsibility and authority are not the same. Even as learning to shoot does not give one license to fire the weapon at will, having responsibility at the "root" level to access and change data and machine operations does not authorize one to do so in all cases and without further sanctions.

With this in mind, the project team asked respondents how many steps (machine or person actions) would be required to get from the data collected in their system logs to the actual identification of an individual. Essentially, the research team wanted to know how close identification was to the information yielded by system logs.

Steps to Identification

Analysis of the responses of system administrators for 71 systems resulted in a mean of 1.6 steps from log data to the identification of an individual. Steps were described as follows: "I would take the IP address from the log data and look up the individual's name to whom the IP was assigned in the IP database." This counted as one step. Another respondent said, "I would take the user ID and go to the accounts database where user ID and personal name are matched." This also counted as one step. If the respondent had said, "I would take the user ID, ask for permission to access the accounts database, and then match the user ID with the person's name," this would constitute two steps. Note that FERPA protections apply if the record is "personally identifiable to a student." It does not require a student's name for identification. FERPA protections do not require a name, but this project took identification to that level.

Only five individuals in the entire sample of respondents indicated that they could not get to a personal identification from the information collected in their logs. The vast majority of respondents indicated that getting to the actual name of a user would be a trivial exercise. Most also indicated that they already have sufficient authority to look at any number of databases that, when accessed, would provide personal identifiers of many types. Sixty-nine (69) percent of those whose logs yielded IP addresses and whose logs yielded domain names were within one step of individual names. Seventy-two (72) percent of those whose logs yielded date and time stamps were within one step of individual names. Eighty-one (81) percent of those whose logs yielded user IDs were within one step of individual names.

If the number of steps to identification were related in some way to the particular category of responsibility held by a system administrator, *e.g.*, central, network, administrative, then researchers could expect to see a correlation in the data analyses between these two elements. They might expect that those individuals whose job it is to provide system administration on large databases and administrative systems containing student registration and admissions data, for example, by virtue of the level of data sensitivity in those systems, would require more steps to get from log data to individual identifiers. Likewise, if the number of steps to identification were related in some way to the primary functions of the computer systems being described, *e.g.*, servers, network monitors, mainframes, then researchers could expect a correlation between these two elements. For example, they might expect that those systems that perform as login servers transporting critical information regarding identification, authentication, and authorization to key services on the campus, by virtue of the sensitivity of data there, will require more steps to get from log data to individual identities.

An adjusted chi square analysis was done on these paired elements of study data, respectively. The analysis showed no significant correlation in either of these pairings. Researchers must conclude that the number of steps required to get from log data to the identification of an individual is simply a function of access to the machine, skill and knowledge of the system, and the availability of the data. At present, it has little or nothing to do with the levels of sensitivity of the data on the systems or with the requirements of confidentiality that users might place on their data. Levels of authorization

and checks and balances to the "authorizations" that currently are affiliated with role and that are assumed to be granted must be put in place to protect personnel and to secure systems.

Log Analyses

Data increasingly are collected in logs. Are they being analyzed and used? Given the time constraints already described by system administrators, the LAMP team wondered how much these data were actually being used. The team asked respondents how often they analyzed the data they collect. Data that are collected but not used and that are stored for undetermined periods of time (and perhaps inadequately protected) increase institutions' potential liability if those data are sensitive and protected by law.

Table 6 shows that approximately 41 percent of system administrators indicated that they analyze the logs they collect daily. However, 41 percent also reported that they analyze

TABLE 6: FREQUENCY OF LOG ANALYSES

Used	Frequency	Percent	Valid Percent
Valid			
Very infrequently or never	5	6.4	6.6
Case-by-case	26	33.3	34.2
Daily	31	39.7	40.8
Weekly	8	10.3	10.5
Monthly	3	3.8	3.9
Other	3	3.8	3.9
Total	76	97.4	100
Missing			
System	2	2.6	
Total	78	100	

the logs very rarely, never, or only on a case-by-case basis when needed. Logs collected but rarely or never analyzed introduce problems of security and data protection discussed below (see the archiving section of this report). Approximately 11 percent analyze their logs weekly. For those who log for the purpose of security (*e.g.*, catching ongoing attacks), for operations management (*e.g.*, to discover and debug software problems), or for network maintenance (*e.g.*, catching login errors and traffic overloads), being able to analyze logs only on a weekly basis is not adequate. Under such conditions, administrators cannot be proactive, but only reactive, and may even be left with nothing to do but clean up the damage after an incident has occurred.

Of those administrators who analyze their log data primarily by looking for patterns that would indicate systems management problems, such as error messages, activity overload, unexpected changes to files, usage level peaks, failed transfers, and so on, there is a nearly even split between those who do so on a case-by-case basis and those who do so on a daily basis. Those who analyze their logs when notified of an event account for 41 percent of respondents; those who analyze logs daily account for 32 percent. System administrators who analyze their log data primarily by looking for security abuse patterns such as multiple attempted logins, unauthorized access attempts, significant command sequences, scanning attempts, and so on are more likely to do so daily than on a case-by-case basis (47 percent daily versus 37 percent when notified of an incident).

As system administrators attempt to secure their systems and proactively search for trouble signs by analyzing log data, they identify patterns of user behaviors that signal probable misuse or abuse of resources. For example, multiple logins on one account within a short period of time from different and geographically distant locations likely signal misuse of an account—usually multiple people using the same, probably stolen or given away, password for account access. Operating rules can be established internal to the system and can be used to signal trouble on data stores and on administrative data files, *e.g.*, cryptographic checksums and hashing can be used to detect when data have been changed. A reasonable way to increase vigilance is to actively analyze patterns on system logs and establish such mechanisms to alert managers to potential problems. But at what point do data collection and analysis become surveillance activity targeted toward individuals?

If identification of a specific individual generally takes less than two steps—and only one from log data already collected—are other data collections and analyses of behavioral patterns happening that might constitute a higher level of logging and perhaps surveillance? Do the data collected in logs ever begin with the identity of an individual and then aggregate more data about that person? The project team asked, "Do you ever target specific activities/patterns of individuals and then collect logs when those characteristics occur?"

Forty-two (42) percent of respondents indicated that on occasion, they have collected information regarding the online activities of particular individuals on their campuses. At times, such individuals were suspected of abuse or misuse of systems and were being investigated. At other times,

particular individuals were identified as being suspect and therefore were watched for potential wrongdoing. In a small number of incidents, targeted information gathering was the result of a specific request from a higher authority seeking to gain information about specific individuals' online behaviors. (See Section 6: Conclusion and Recommendations for further discussion of targeted personal information gathering.)

Archiving Log Data and Why

The terms default, enabled, and customized/scripted were seen to have caused confusion earlier in this study. In the investigation of archiving, the terms archive and back-up also seemed to cause confusion for respondents. During the interviews, some system administrators used the terms interchangeably. Although it is possible that the terms are used consistently within their organizations, we urge caution when drawing conclusions from responses to the following questions: "Do you archive log information you collect?" and "For how long?"

The LAMP technical focus group provided input as to why project respondents seemed to confuse these terms. The group defined archiving as the action of recording and storing copies of logged data for the purpose of freeing up disk space and for analysis of incident investigation or for a record of the system's operational history. The term back-up refers to the practice of recording the system state and current data for use in case of a system failure. In the event of a system procedural error, a system back-up such as magnetic tape can be used to restore the system to its last known functional state. The focus group noted that these terms' similar meanings result in even the most well trained technical specialists using these words haphazardly and interchangeably. Nevertheless, there is a clear difference between these terms that is not always recognized.

Responses to the question about archiving showed that a high percentage of administrators (82 percent) do archive log data. They described two basic reasons for doing so: (1) for history and retrospective analysis (It is important, especially in the case of security incidents that may have to be investigated, that data be available for analysis.); and (2) to maintain a steady state on the systems they manage. Many referred to this as back-up so that systems could be reconfigured and understood in the event of systems failure. A majority (65 percent) of respondents maintained their archived data for more than two months; 22 percent maintained archives for between one and two months. The research team did not ask about the longest period of time for which archives might be kept.

Access, Criteria, and Policies

If the data on the machines are very to extremely sensitive; if the logs themselves contain information such as IP address, user ID, and other information that is generally only one step away from identification of individual persons; and if the majority of log data at colleges and universities is being archived for greater than two months, then how, if at all, are these data being protected? The research team asked if authorizations were required to access these data and, if so, what the access criteria are. They also asked if policies existed to guide the community and technical staff in appropriate use of, and access to, these data.

Response to the question about authorization to access log data was strong. Virtually every system administrator (97 percent) indicated that authorization for such access is required. In most cases, however (84 percent), authorization is synonymous with role. Individuals who were assigned to a particular work group, who were at a particular level of system administration, who had "root" privileges on the major systems being described, were also felt to have "authorization" to access the log data simply by virtue of their role,. Twelve (12) percent required a second level of approval—permission of the designated manager—to access the data. Note that responsibility and authority are not the same. Note, too, that at most colleges and universities, individuals with considerably less experience in the field than our respondents, and often, students, are given system administrator responsibilities with root access to systems.

TABLE 7: NUMBERS OF INDIVIDUALS WHO HAVE ACCESS TO LOG FILES

Number who have access	Frequency	Percent	Valid Percent
Valid			
Fewer than 5 people	42	53.8	56
5-10 people	21	26.9	28
11-25 people	10	12.8	13.3
More than 25 people	2	2.6	2.7
Total	75	96.2	100
Missing	**99**	**1**	**1.3**
System	2	2.6	
Total	3	3.8	
Total	78	100	

Because the size of departments varies greatly, the number of people who are in roles granted "authority" to access these data also varies. Table 7 shows the percentage of respondents that selected each of the categories designating the numbers of individuals who have access to log files.

The fact that system administrators often change roles and responsibilities (as shown above in the length of time they reported being in their current jobs) increases concern about this method of granting "authority." It is important to understand what precautions are taken to remove access privileges as individuals vacate these roles. Likewise, as new people are assigned to these roles, it is important to understand the amount of training they receive in the protection and management of sensitive data, institutional policies, and relevant law. Policies are helpful in guiding new as well as experienced employees. However, 65 percent of respondents reported that they did not have written policies regarding log data, its collection, its appropriate use, authorizations for access, or disposal of such data. If unwritten practices alone are depended upon, then practices are sure to change as people change roles. This is not an acceptable management practice for systems handling sensitive data.

Data Analyses Results

Logging and Its Relationship to Student Records

This section explores the question of whether the law that guides the handling of student records would be helpful in informing the handling of log data as well if such data constitute education records. To answer this question, the LAMP study team solicited college and university registrars' expert opinion as to whether the information obtained through logging on college and university campuses constitutes a "record" under the Family Educational Rights and Privacy Act (FERPA). Designated the stewards of student data, these individuals have the authority, in conjunction with university counsel, to interpret FERPA and to put procedures in place that will protect student records at their respective institutions. Fourteen of the 16 registrars agreed to participate in the LAMP project.

The project team developed nine scenarios on the basis of real-life incidents. The team and the project's advisory board reviewed the scenarios. Six were selected[9] and mailed to the participating registrars. Each scenario was followed by five questions concerning "education records," appropriate educational uses, violations of FERPA regarding gathering or sharing such data, and access as "legitimate educational interest." Participants were asked to review the scenarios and to respond over the phone to each of the questions. (See Appendix F for the Registrar Scenarios and Data Collection Instrument.)

In addition to collecting data from participating registrars, the team asked two individuals from the Department of Education's Family Policy Compliance Office (FPCO) to review the scenarios and respond to the questions. (These individuals are referred to hereafter as "the experts.") The team anticipated that the experts and the participating registrars would supply the insight needed to understand the status of log data and the appropriate use thereof. Basic descriptive statistics provided information about agreements between registrars and the experts and about different types and levels of log data.

The participating registrars and the experts from the Family Privacy Compliance Office were asked the following five questions for each of the six scenarios presented:

QUESTION 1: Does any of the information referred to in this scenario constitute an education record under FERPA?

QUESTION 2: Does collection of this information constitute a violation of FERPA?

QUESTION 3: Does the sharing of this information constitute a violation of FERPA?

QUESTION 4: Is this an "appropriate educational use" of student information under FERPA?

QUESTION 5: Does access to this information qualify under "legitimate educational interest?"

Table 8 presents data pertaining to the six scenarios—five questions about each—and the registrars' and experts' responses.

The Scenarios and the Results of Analyses

Scenario I: Let Me Know

John, a system administrator in a college math department, collects information on the operation of the department's computers and networks. He needs to watch the flow of traffic over his systems to responsibly manage the electronic resources. Students, faculty, and staff heavily use the computers. John's log information includes the number of people who sign on, the number of people using mail at any given time, the number of packets of information sent across the networks, and the static IP addresses of the machines from which, and to which, those packets are sent. Generally, John does not collect information about the kinds of services individual users seek, where they go on the World Wide Web, how much time they spend emailing or downloading files, or specifically what files they look at. However, today is different. The Dean wants to know how three student staff members are using their time. He has demanded that information be collected from their machines about them and their use patterns and that it be turned over to him for analysis. John secretly turns on the logging function and collects information from the machines used by these students.

9 The six scenarios were selected because they represented a range of different factors that might be relevant to decisions about record status, legal behavior, intended use of records, etc. Each represents a different type of incident, a different amount of data collection, and a different level of potential intrusion into privacy.

TABLE 8: PERCENT RESPONSES TO SCENARIOS BY REGISTRARS AND EXPERTS

	Scenario I:	Scenario II:	Scenario III:	Scenario IV:	Scenario V:	Scenario VI:
Qustions	**Let Me Know**	**24x7 Information**	**Complaint Follow-up**	**Campus Safety...**	**Watch This One**	**Printer Server Log**
Question 1						
No	36	14	36	0	36	21
Yes	64	86.0*	64.0*	100.0*	64.0*	71.0*
Depends	0	0	0	0	0	0
Don't Know	0.0*	0	0	0	0	7
% Matches**	0	86	64	100	64	71
Question 2						
No	86.0*	93	79.0*	79.0*	79.0*	100.0*
Yes	14	7	21	14	21	0
Depends	0	0	0	7	0	0
Don't Know	0	0	0	0	0	0
% Matches**	86	93	79	79	79	100
Question 3						
No	50	43	50	21	29	64
Yes	36	21	36	43	57	7
Depends	7	14	7	7	7	7
Don't Know	7.0*	21.0*	7.0*	29.0*	7.0*	21.0*
% Matches**	7	21	7	29	7	21
Question 4						
No	64	14	50	7	57	21
Yes	14	57	36	57	29	64
Depends	14.0*	14.0*	14.0*	14.0*	14.0*	14.0*
Don't Know	7	14	0	21	0	0
% Matches**	14	14	14	14	14	14
Question 5						
No	57	21	57	14	71	21
Yes	29.0*	64	36	64	21	71
Depends	7	7.0*	7.0*	7	7	0.0*
Don't Know	7	7	0	14.0*	0.0*	7
% Matches**	29	7	7	14	0	0

*** "Percent matches" means the percentage of the time the experts' responses matched those of the registrars for that question.*

Note: Percentages may not total 100 because of rounding.

The information includes time of access, the specific services to which the student staff members go, what content they are viewing, how long they are at particular sites, what email and to whom email is sent, and when they log off their machines.

The project staff felt that several aspects of this scenario were significant: The scenario is about students yet the students are also employees of the institution. The person requesting the data is a dean, with authority over the system administrator being asked to collect data pertaining to the behavior of the student employees. The request is for secret logging, without the student employees being notified. The request for log information goes beyond machine data to personally identifiable information about electronic mail, web, and file use. The scenario illustrates low-level surveillance.

Sixty-four percent of respondents agreed that the data gathered in the logs referred to in Scenario I did constitute an education record; however, there was 0% matching with experts on this item. In fact, the experts replied "don't know," perhaps signaling the need for additional information. Eighty-six percent of registrars agreed that collection of these data did not violate FERPA. The percentage agreement with the experts on this item was high. Fifty (50) percent felt that the sharing of these data did not violate FERPA, while 36 percent said it did. A majority of registrars thought this was not an appropriate educational use, and a small majority said this did not qualify as legitimate educational interest. Matches between registrars and the experts on these items were low—14 and 29 percent, respectively.

Scenario II: 24x7 Information

Tudlo College is electronically advanced. Its staff members are particularly proud of the fact that all of their computer resources have been networked together using state-of-the-art technology. If a staff member needs information regarding a particular student, she can obtain it quickly and efficiently. The information the institution stores is extensive: Not only does it maintain student grades, addresses, course selection lists, financial aid, and other important information online, but it has begun to collect information about student transactions, as well. Through extensive networks and the use of card swipe machines and key card access mechanisms across campus, information regarding student transactions is collected twenty-four hours each day, seven days per week. Staff thus can learn when a student leaves his residence hall, which building he enters and when, if he enters a particular room in that building, and even when he leaves. Information about if and when the student makes a purchase, eats a meal, reserves a book at the library—even where he parks his car and how long he stays in a particular parking space-can be gathered. Each time the student swipes his ID card, data are collected and stored in the college's databases. In the event of an emergency, the college can locate a student rapidly and with great precision. College officials believe that student safety will be enhanced by this new capability.

Project staff believed the following aspects of Scenario II were significant: The incident involves a student. The information collected pertains to grades, financial aid, addresses, course enrollment, financial transactions, locations, books borrowed, meals taken, and dining times. The information that is available is from multiple databases networked together. Depending on the use made of it, such a collection of data can provide extensive information about an individual and may constitute a high level of surveillance

Registrars strongly agreed (86 percent) that the data referred to in Scenario II did constitute an education record according to FERPA. There was an 86 percent match with the experts on this question. There also was very strong agreement (93 percent) that collecting such information would not violate FERPA; the percentage match with the experts was equally high. However, there was considerable confusion and opinions were mixed about sharing the data. Respondents barely agreed that this was an appropriate educational use. Sixty-four percent agreed that there was a legitimate educational interest in accessing such information.

Scenario III: Complaint Follow-up

Fred is a system administrator for a very large college in a midwestern university. He receives a complaint from one of the college's sophomores. The student, Sarah, describes a series of email messages that she has received from a person whose name she does not know. Though the name on the email is not familiar, the pattern of communication is. Sarah feels certain that she knows who is sending her these repeated messages. She alleges that Kenny, a senior at the college, is stalking her. She wants Fred to take action as soon as possible. Sarah reports that even when told to stop sending her email, the sender did not stop and in fact increased the demands in his messages. The sender tells her what time her first class meets, which building she goes to in the morning, and where and when she eats meals. Sarah is frightened and consequently is having difficulty studying. From system logs, Fred can identify the network, the machine, and even the account from which the email messages were sent. He is certain that the account is in Kenny's name. Sarah wants action. Fred has some information that points to Kenny. He decides to follow up by intensifying the collection of information on Kenny's machines. He writes a script that will alert him each time his account is in use and that will provide information on the account. In addition, he establishes a system with the residence hall advisors to tell him when Kenny is in his room and therefore using the IP address to which his machine is assigned. He soon will have a large quantity of information about Kenny's electronic activities. His plan is to watch Kenny's online behavior, analyzing log records for at least two months.

The team believed the following aspects of this scenario were significant: that the information is about two students; it involves a charge of harassment by one student against

another but no evidence thereof; it is secret investigation and data gathering regarding one of the students by the system administrator who is an employee of the university.

Sixty-four percent of the registrars agreed that these data do constitute an education record. Seventy-nine percent agreed that collection of these data does not violate FERPA; the percentage match with the experts was equally strong (79 percent). Agreement that sharing these data does not constitute a violation of FERPA was weak, and opinions were mixed. Agreement that this is not an appropriate educational use was similarly weak (50 percent vs. 36% who think it is an appropriate educational use). Fifty-seven percent of respondents indicated that access does not represent a legitimate educational interest whereas 36 percent indicated their belief that it does.)

Scenario IV: Campus Safety Needs It

As the university becomes more electronically connected, Sergeant Denver, an officer with the Department of Campus Safety, is delighted with the information that potentially will be available. Students are asked to have their pictures digitized for inclusion on ID cards. They use keycard access for entry and departure from buildings, their unique names and machine IP addresses are stored in readily accessible university databases, and their course and meal-time schedules are made available to the campus safety staff as well. Sgt. Denver particularly favors the availability of pictures online. This is information that Campus Safety needs. Through the use of cameras, these images can be matched with individuals entering campus buildings, or those stopped for suspicious behavior, and their identities can be verified. If they do not belong on campus, they can be readily identified as non-students. If they are students but are in locations without authorization to be there, they can be notified or reprimanded. Witnesses to crimes can even be shown the pictures so they can identify suspects in "virtual lineups," thereby assisting the officers in their work. The picture information can be connected to applications for employment or to resumes being sent to graduate or professional schools. College departments already are using them to assist professors in identifying students in large classes. Keeping this information online facilitates access and rapid transit when it is needed by different organizations on campus and officers in their different locations. It can be kept available indefinitely.

The project team believed several factors within Scenario IV to be interesting and significant: The scenario involves students. The university collects and stores copies of student images as digitized photos. There are a number of appropriate ways in which faculty, students, and staff might use images of students. Though the images were taken explicitly for use on personal campus identification cards, campus police wanted to make other uses of these same images. Police wanted to combine these stored images with other stored electronic data for investigation of crimes and potential identification of criminals. This scenario potentially involves an unauthorized secondary use of personal data.

Scenario IV drew unanimous agreement (100 percent) that the data do constitute an education record according to FERPA, and the percentage match with the experts was 100. There also was a high percentage of agreement (79 percent) that collection of such data does not violate FERPA. (The percentage match on this item was 79. There was confusion and disagreement about whether sharing this information violated FERPA. There was slight agreement that this situation does represent an appropriate educational use, and 64 percent agreed that this also represents a legitimate educational interest (only 14 percent said it does not.) For both of these items-educational use and legitimate educational interest-the percentage match with the experts was only 14.

Scenario V: Better Watch This One

Sidney is in charge of all network systems at a small liberal arts college. He also manages all of the central services and servers for the college. It is very difficult to keep everything up-to-date, working together, and coordinated. Sidney is not happy with the number of interruptions caused by unknown hackers who seem to explore systems looking for unprotected files and applications. He has configured his machines to log all machine activities, the amount of traffic flowing over the network, the number of machine errors that occur, and to notify him and other system administrators when machine problems occur. But this information does not communicate much about individual computing behaviors. Sidney knows that several students on campus have extraordinary computer skills and pride themselves on being able to do remarkable electronic feats on the networks. While they have never been identified as causing a computer problem or disciplined for inappropriate behavior on campus, Sidney decides that it is better to periodically watch the students than to remain ignorant of their activities. He selects one in particular who he thinks would be likely to engage in hacker activity and intensifies the information gathering logs connected with the student's ID, unique name, account use, the IP address of his machine,

and all connections made from his residence hall. By periodically gathering and analyzing this extensive information about the student's local and web activity throughout the year, Sidney can be assured that the student is not participating in hacker activity—or, if he is, he can be quick to suspend his account or report him to Student Affairs for discipline.

The following factors in Scenario V were identified by the project team as significant: The scenario involves one selected student; there has been no violation of university policy or law; there is no evidence of wrongdoing on the part of the student; information about the students is collected secretly; the information monitored includes Web activity, IP addresses, and much more. This scenario represents a substantial form of surveillance.

Sixty-four percent of respondents agreed that the data referred to in Scenario V do constitute an education record under FERPA. There was even stronger agreement (79 percent) that collection of these data does not violate FERPA. The percentage match with the experts was equally strong (79). There was slight agreement (57 percent) that sharing would constitute a violation of FERPA and that this scenario did not represent an appropriate educational use (57 percent). Unlike the other scenarios, about which opinion regarding legitimate educational interest was mixed, there was relatively strong agreement (71 percent) that access to these data did not constitute a legitimate educational interest. The percentage match with the experts on this item was zero because the experts said they "didn't know," perhaps signaling the need for additional information.

Scenario VI: Printer Server Logs

Administrators at a central university are becoming alarmed at the cost of providing computing resources to the campus community. They are struggling to find ways to pass the costs of particular services back to the students and in some way to "throttle" the overuse of resources they suspect. To do this, they must gain access to name-related use patterns on specified services. System administrators are instructed to initiate logging mechanisms on all print servers in the public computing sites. The data logs provide the following information: the name of the student printing material, his or her unique name and account number, the number of sheets of paper printed, the time and date of the printing, whether the printing was graphics or text type, and the machine from which the printing request was sent. The logs are analyzed, and a bill for printing services is sent to each student. Logs are kept for two to four years for purposes of accounting.

Project staff identified the following factors in Scenario VI as significant: The situation involves students. The data collected are personally identifiable, including name, account number, and quantity of printing. The purpose of the logs is to bill students for printing charges. Data are kept for two to four years. The data represent business processes directly connected with the university and students' business relationship with the university.

Scenario VI drew strong agreement among the registrars. Seventy-one percent agreed that the data do constitute an education record under FERPA; the percentage match with the experts was 71. Registrars were unanimous (100 percent) in their belief that collection of these data did not constitute a violation of FERPA; the percentage match with the experts was 100. Sixty-four percent agreed that sharing these data did not constitute a violation of FERPA. Sixty-four percent also agreed that what was described probably was an appropriate educational use. Finally, 71 percent of the registrars agreed that access to this information did constitute an appropriate educational interest. The percentage match with the experts on this item was zero because the experts replied "It depends," likely signaling the need for more information.

Summary—Responses to Scenarios

In summary, the six scenarios presented to participating registrars and the experts pointed out several areas of strong agreement, as well as several areas of confusion and uncertainty. There is strong agreement that the data collected in each of the six scenarios constitute an "education record" under FERPA. This agreement was particularly strong for scenarios II (24x7 data collection and IV (student online pictures). Agreement between registrars and the experts was highest for these two scenarios—86 percent and 100 percent, respectively.

There is strong agreement that collection of the data represented in the six scenarios does not constitute a violation of FERPA. Particularly strong agreement was evident for Scenarios I (student employee logging), II (24x7 student data collection), and VI (billing for printing). Percentage match with the experts was strongest for these three scenarios as well, at 86, 93, and 100, respectively.

All of the scenarios point out confusion and uncertainty regarding whether sharing of the logged data constitutes a violation under FERPA. Possible reasons for this may be (1) the respondents had not previously been faced with questions

about these kinds of data; (2) the law is unclear ; (3) the question was poorly worded; or (4) more information was needed. However, had this question required more specific information, researchers would have expected a higher percentage of "it depends" responses. Instead, only 7 to 14 percent of respondents replied "it depends." (See Section 6: Conclusions and Recommendations for further discussion of this response.) Across all of the scenarios, the highest percentage of matches on this item between the registrars and the experts was 29, suggesting that little guidance is available on how to answer this question or that no guidelines regarding this type of data sharing currently exist.

Responses as to whether the scenarios described "appropriate educational use" of student data were similarly mixed. While the percentages of agreement among registrars were not as low as those recorded in response to the item on the sharing of data, the highest percentage of agreement on this item was only 64. With regard to Scenario I, 64 percent of respondents said that collecting information by secretly logging student employee computer usage was not an appropriate educational use of data; with regard to Scenario VI, 64 percent said that logging printer usage for billing purposes was an appropriate educational use.) Again, the percentage match between registrars and experts for all scenarios was very low.) Many more respondents indicated that it depended on other factors. The question may not have been well worded, and the terminology itself may be unfamiliar or confusing. It also is possible that other specific pieces of information that are spelled out under the law should have been included in this scenario in order to clarify respondents' choices.

Respondents were more certain about whether access to the data described in the scenarios constituted a legitimate educational interest. This was particularly true for Scenarios V (watching potential hackers) and VI (billing for printing). With regard to Scenario V, respondents agreed that monitoring potential system abusers was not a legitimate educational interest whereas billing for printing (Scenario VI) was a legitimate educational interest. Across all of the scenarios for this item, the highest percentage match between registrars and experts was 29 (Scenario I, logging student employee activity), indicating that the experts needed more information about how individual institutions have defined "legitimate educational use" to answer the question. More specific guidelines on this item may also be needed.

To determine why responses about data sharing, educational use, and legitimate educational interest differed, we must refer to the language in FERPA and interpretations thereof.

The Family Educational Rights and Privacy Act

It is not the intent of the LAMP project team to attempt to interpret FERPA, to offer legal advice about the status of records, or to judge the appropriateness of disclosure of records under the law. However, it is appropriate to comment on the results of the data analysis relative to language in the Act and commonly accepted requirements of the Act. These comments follow the order of the five questions asked in the registrar data collection instrument.

QUESTION #1

The first question pertained to whether the data collected in logs, as described in each of the scenarios, constituted an education record as defined by FERPA. Researchers learned that there was strong agreement among respondents across all of the scenarios that the data were indeed education records. The Act[10] clearly defines education record; therefore, identifying the appropriate response apparently was relatively straightforward.

"Education record means those records, files, documents, and other materials which contain information directly related to a student and are maintained by an educational agency or institution or by a person acting for such agency or institution.[11]

"Record is understood to mean any information or data recorded in any medium (*e.g.*, handwriting, print, tapes, film, microfilm, microfiche, any form of electronic data storage."

The log information described in each of the scenarios was individually identifiable: directly related to a student and retained by the agency/institution. A majority of the respondents judged it to be an education record in accordance with FERPA. However, there was little match with the experts on this item. This mostly likely was the result of some question as to whether the student was employed "as a result of his or her status as a student." The law states that records that are kept as part of the normal course of business are not education records unless the records are for a student employed "as a result of status as a student," in which case they are education records. In Scenario I, the

10 AACRAO, 1995, "Guidelines for Postsecondary Institutions for Implementation of the Family Educational Rights and Privacy Act of 1974 as Amended."

11 Rainsberger, R., E. Baker, D. Hicks, B. Myers, J. Noe, & F. Weese. 2001 "The AACRAO 2001 FERPA Guide," AACRAO Professional Development & Education Series, The American Association of Collegiate Registrars and Admissions Officers. Washington, DC.

difference of opinion among respondents and certainly between the registrars and the experts on the "records" question most likely hinged on the student's employment status.

QUESTION #2

The second question respondents were asked was whether collection of these log data would be considered a violation of FERPA. Respondents strongly agreed across all six scenarios that collection of the information described did not constitute a violation of FERPA. Careful reading of the Act shows that the law does not address the particular types of information institutions may collect regarding students. The Act is silent on this point, and respondents' strong agreement seemed to reflect the lack of regulations in this area.

QUESTION #3

The research team asked if sharing of the log information constituted a violation of FERPA. Data analysis revealed confusion and uncertainty in the registrars' responses to this question across the six scenarios, with some differentiation in specific cases. The word "sharing" does not occur in FERPA. However, if you share information, you disclose it by default. The relevant question, then, is with whom you share it. "Disclosure" is used and defined as follows: "To permit access to or the release, transfer, or other communication of personally identifiable information contained in education records to any party, by any means, including oral, written, or electronic means." In only two of the scenarios were there actions that implied that "disclosure" did or would take place. The question was based not on actual wording within the scenarios but rather on the concept and whether sharing of the information that was described would constitute a violation of FERPA.

There are several potential reasons for the difficulty respondents seemed to have in answering this question: (1) They may not have had experience with disclosure of the types of data—log data—described in the scenarios. (2) There are no clear guidelines relative to log data in particular. However, because such data are personally identifiable and qualify as education records, this should not have been the primary source of confusion. (3) The differing responses to this question most likely stemmed from the fact that the law allows institutions to define for themselves to whom they will disclose information, albeit it must be stated in policy and made readily available to students. For example, institutions may define who is a "school official"-those who may legitimately access student information. Institutions also may define the criteria for determining what constitutes a legitimate educational interest. There is no evidence in Scenario III that they have done either.

While institutions may define processes for sharing information, the intent of the law seems clear: to protect the privacy rights of students for information that is not designated as public/directory information and to give parents and students the right to determine which items of their information contained in the institution's directory information may be disclosed. The law states, "Directory information means information contained in an education record of a student which would not generally be considered harmful or a violation of privacy if disclosed."

Answering the question about disclosure/sharing of the log information therefore depended on whether the institution had fulfilled its annual notification requirement, whether it had designated the intended recipient of the shared data as a "school official" according to their job description, and whether the intended access to the data fell within the institution's criteria for what constitutes legitimate educational interest. None of the six scenarios provided sufficient information about these aspects of institutional policy and process to allow respondents to derive an unequivocal answer to this question.

QUESTION #4

The fourth question that was asked for each of the scenarios was "Does this constitute an appropriate educational use" of the data? Certain commonly used terms in the implementation of FERPA have gained acceptance over the years. Legitimate educational interest, the terminology used in the law, has come to be known more familiarly as "need to know." The term appropriate educational use is not defined in the law, and while it may be understood, it is not necessarily accepted as a substitute. Respondents may have had personal opinions about whether the data use described in the scenarios was "appropriate, since there is no legal definition," and their responses reflect the different cultures of the institutions represented and their particular sets of values. The variance in the results may indicate several issues pertaining to "appropriate educational uses of data" that need to be discussed and decided, particularly as we now have the capability to transport and disclose information electronically.

It is important to note that the law allows disclosure of education records without prior student consent in certain cases—specifically, disclosure to state and local officials or authorities to whom the information:

is specifically allowed to be reported or disclosed pursuant to state statute adopted before November 19, 1974, if the allowed reporting or disclosure concerns the juvenile justice system and the system's ability to effectively serve the student whose records are released…[12]

Scenario IV involved campus police requesting the release of information for their use in identifying potential criminals. Because campus police wanted to use information about students for whom there was no suspicion of wrongdoing "for virtual lineups" or image comparisons, respondents were faced with a serious question of appropriate use. No subpoenas are indicated in the scenario. Campus police may be defined as employees of the institution and possibly as "school officials." The question in fact may be even more complicated. Release of student education records to campus safety personnel is not an automatic process that is not subject to the standard FERPA protections in regard to disclosure. Simply because the law enforcement personnel may be employed by the institution does not nullify the precautions that need to be taken in making any disclosure. "Law enforcement agencies and health offices requiring access to student records are bound by all of the conditions of accessibility stipulated by FERPA."[13] Further, the records identified in Scenario IV cannot be assumed to be exempt from the definition of education record for law enforcement purposes. That exemption is for "records maintained by a law enforcement unit…created by that law enforcement unit for the purpose of law enforcement." Photographs of students qualify as an education record. They were not created for the purpose of law enforcement. In recent revisions of the law, institutions have been given the right to define photographs as directory information if they wish to do so. Given the many inappropriate ways in which photographs can be used, as well as student concerns about pre-employment and pre-admission discrimination and even potential stalking through the use of publicly available photographs, institutions need to carefully consider use of these records. Certainly this is an area needing further clarification and guidelines.

QUESTION #5

Finally, respondents were asked if the scenarios described access to information that represented a "legitimate educational interest." Although the law does not define this term specifically, it is clear that institutions must do so. "It (law) states that institutions must establish their own criteria, according to their own procedures and requirements, for determining when their school officials have a legitimate educational interest in having access to a student's education records." Even if a person has been designated a "school official," he or she does not have inherent rights to any and all education record information. The school official must demonstrate a legitimate educational interest as opposed to a personal or private interest, and such a determination must be made on a case-by-case basis."[14] The law states that the institution must publish its policy and criteria for designating school officials and determining legitimate educational interests and make them available to students. It further states that when disclosures are desired that have not had prior notice, decisions must be made on a case-by-case basis (unless they meet designated exceptions); this was not evidenced in any of the scenarios.

"Guidelines for Postsecondary Institutions for Implementation of the Family Educational Rights and Privacy Act of 1974 as Amended," a publication of the American Association of Collegiate Registrars and Admissions Officers (AACRAO), provides sample criteria for determining who are "school officials." Criteria include a person employed by the institution in an administrative, supervisory, academic, or research or support staff position. If an agency had listed such criteria and had designated system administrators as school officials, then registrars could have answered "yes" to at least part of the question about "legitimate educational use." However, answers still would hinge on whether the institution's criteria for assessing legitimate educational use included the use implied in each of the six scenarios. Without this information about agency/institutional policies and criteria, registrars had to answer "It depends" or provide a response based on their own personal values. Only in scenarios V (Better Watch This One) and VI (Printer Server Logs) did respondents register strong opinions.

12 Ibid.
13 Ibid.
14 Ibid.

Conclusions and Recommendations

On July 11, 2001, House Majority Leader Dick Armey and the American Civil Liberties Union released a joint statement entitled "Proliferation of Surveillance Devices Threatens Privacy." The statement describes several incidents of technology use for secret surveillance and collection of data regarding United States citizens. Armey and the ACLU write:

We are extremely troubled by this unprecedented expansion in high-tech surveillance in the United States. We believe that technology should not be used to create a "virtual line up" of Americans who are not suspected of having done anything wrong. The threats to privacy in America are all too real. We believe the privacy risk outweighs any benefits that these devices may offer. It's time to take notice of what has happened to privacy in America today.[15]

One day earlier, on July 10, computer security experts and individuals from many countries and whom the public sometimes calls hackers were convening to learn more about new security issues facing companies and education institutions that are heavily dependent on information technologies and networks. The CERT Coordination Center, a security emergency response and research group at Carnegie Mellon University, reported that more than 7,000 computer security violations had been reported in the first quarter of the year-a number much greater than in previous years. Kevin Manson, a conference attendee, said, "I'd say we're about as well prepared for cyber warfare as we were the day before Pearl Harbor was struck."[16] William Tafoya, a computer security expert, stated, "With the current wave of computer crime, it's not a matter of having enough people, it's a matter of not having the expertise."[17]

The LAMP project was initially conceived precisely because college and university policy makers were not prepared to define the relationship between log data and records. Direction and even guidance regarding the relationship were not available. Experts who could relate what was being done in terms of logging and monitoring activities on campuses were not knowledgeable about the legal protections of privacy afforded under FERPA, and vice versa. With funding support from the National Science Foundation, Digital Government Program, the project was undertaken to answer this and related questions.

Summary

The LAMP project provides information about a number of specific issues, *e.g.*, how much logging is being done, how easy it is to access personally identifiable information. The project also sheds light on more general issues, such as the trend to use technology without fully understanding its capability, the creep and expansion of surveillance potential through technology, and the need for more extensive interdisciplinary communications and for discussion of ethics as well as the law.

Sixteen geographically clustered colleges and universities participated in this study. The institutions varied in size, in their experience in the use of electronic networks, and in their missions (public vs. private, two-year vs. four-year). The first group of participants was system administrators. Fifty-seven system administrators participated in the study-an average of approximately three from each institution. On each campus, system administrators who were considered the most knowledgeable about computer logging—"those who know the most about computer logging on networks, administrative systems, central systems, and for large colleges"—were identified. A member of the project team interviewed each system administrator individually. In response to a standard questionnaire, system administrators provided data regarding their major systems.

The second group of participants was registrars and FERPA experts from the Family Records Compliance Office. Fourteen of the 16 institutions' registrars participated. (Registrars were selected because they are responsible, by virtue of their position, for implementing FERPA on campus and because they possess the knowledge to interpret FERPA regulations). Responding to a selected set of six scenarios describing various logging events and subsequent data use, they provided information about privacy and protections afforded to education records. Two experts from the FRCO office also agreed to review and comment on the scenarios, thereby adding their expertise to this study.

15 "Proliferation of Surveillance Devices Threatens Privacy," Joint Statement issued 11 July 2001. Washington, DC: American Civil Liberties Union, 2001.
16 B. Keefe, 2001. "Hackers Convene amid Signs Computer Security is Eroding," Newsfactor Network, http://www.newsfactor.com/perl/story/11885.html.
17 Ibid.

Results: Specific and Obvious Conclusions

There are specific and obvious, as well as more subtle, conclusions to be drawn from this study.

REGARDING PARTICIPANTS, THE STUDY SHOWS THAT:

- System administrators, the most knowledgeable about logging on campus, had a wide range of educational training; their highest earned degree ranged from a high school diploma to a Ph.D. The majority held science-related degrees (B.S. or M.S.).
- The project system administrators possessed an impressive number of years of experience in the field of computing (71 percent between 14 and 42 years.) However, there was evidence of job changing and a high rate of personnel changes within the population. Nearly 50 percent had been in their current jobs for less than five years.
- Although many of the respondents had participated in extensive self-teaching, they had received limited formal training in security on the machines they administered; less training in data protection and fair information practice; and little or no training in student record protection under FERPA. None had taken courses or seminars concerning FERPA.
- The registrar participants had varying degrees of knowledge about technical logging and monitoring activities.

REGARDING CAMPUS PRACTICES, THE STUDY SHOWS THAT:

- System administrators have a complicated set of responsibilities and operate within considerable constraints of time, resources, and support for training-and this despite increasingly complicated systems.
- Logging does happen on campuses: Ninety-six (96) percent of respondents log on their major machines, and 59 percent of system administrators would like to do more extensive logging. Logging is done for three primary purposes: security; systems and network maintenance/ troubleshooting; and operations management.
- Most of the respondents (72 percent) use the default logging function on the machines they manage. A high percentage (82 percent) disables or enables specific functions to make log output more useful. A high percentage (64 percent) also customizes the logging done on their systems primarily to organize the log output for easier searching, analysis, and use.
- Terminology in this area is used inconsistently. As a result of rapid technology changes, upgrades, and issues inherent in interoperability, administrators appear to be using technology about which they do not have full knowledge.
- Logs typically contain several pieces of information, including IP addresses, user IDs, domain name addresses, and date and time stamps of computer connections and transactions.
- Using the information contained in logs to identify specific individuals typically is a trivial exercise for system administrators, requiring, on average, only one action.
- Though some respondents analyze logs daily, most analyze log data only infrequently.
- Forty-two (42) percent of respondents sometimes use logging to collect information about the behaviors of specific individuals without informing those individuals that they are doing so and without specific permissions being required.
- Many respondents (82 percent) archive log data. Twenty-two (22) percent keep logs for more than two months.
- Logs are archived and typically require special "authorizations" to access. However, those authorizations essentially are tied to role, without additional permissions being required. Most schools do not have in place written policies regarding access to log information.

REGARDING EDUCATION RECORDS, THE STUDY SHOWS THAT:

- The registrar respondents strongly agreed across all of the scenarios that the log data described constitute education records under the definition provided in FERPA.
- Registrar respondents strongly agreed that collection of the log data does not violate FERPA. There was considerable uncertainty about whether sharing such data violates FERPA.
- Responses regarding whether the actions described in the scenarios constitute an appropriate educational use of the data and whether access to such data would be considered "legitimate education interest" under FERPA were mixed.

Results: Dynamic Pressures on System Administrators and Registrars

The LAMP study revealed several interesting trends, as well. Both system administrators and registrars are subject to dynamic pressures brought about by the new electronically networked environments within colleges and universities.

System administrators, educators, and politicians now face pressure for privacy protections on the one hand and for security protections on the other. The need to balance priorities has administrators of computer systems at the pressure point, somewhere between these competing forces. System administrators have demanding, increasingly complicated, and difficult responsibilities. They are expected to perform their duties under conditions characterized by too little time, rapidly changing technologies, increasing attacks on systems, increasing numbers of hardware and software failures and bugs, and rapidly changing personnel. Their responsibilities have increased exponentially. Yet opportunities for training and support to obtain the personnel and equipment they require to maintain and operate their systems have not kept pace. Frequent staffing changes, departmental reorganizations, increasingly knowledgeable but sometimes under-trained personnel, interoperable yet sometimes incompatible systems, and exponentially growing user demand have combined to produce trying work environments for system administrators. These individuals have not had the opportunity to become experts in the intricacies of security on the multiple systems they manage; in the sensitivity characteristics of the data that flow across their machines; or in the privacy considerations and rights of users.

Registrars are the other significant group being asked to balance seemingly contradictory priorities. When the Family Educational Rights and Privacy Act first was written, data administration and processing were accomplished primarily on mainframe computers with tight controls on access and data management. Registrars, as stewards of the data, were more easily able to establish controls and procedures to implement the rules and regulations of the law because there were fewer data access points and more easily coordinated procedures.

Since that time, however, information technology has enabled the distribution of data over electronic networks. Data now can be accessed on the desktops of any number of college and university personnel—students, faculty, and staff. Passing over networks that are yet to be adequately secured, such data also may be available to unauthorized persons through sniffers and other capture devices. The distributed environment has grown exponentially in terms of the number of users and the number of people who believe they have a "need to know"-potential "school officials" under FERPA, if the institution defines them as such. The registrar's responsibility to implement the law in such an environment is onerous. To ensure that policies are in place, when so many departments and individuals may be trying to access data, is a formidable task. To ensure identification of "records," provide annual notification, define the criteria for "legitimate educational interest," etc. in these new electronic environments is difficult and time-consuming.

It is difficult to manage data that can travel, with or without authorization, at high rates of speed, spreading to a multitude of receivers in mere seconds, and to do so in accordance with law and with sensitivity to individual preferences. Registrars have this responsibility. At present, they also find themselves required to be involved in, as well as to understand the complexity of, new enterprise-wide systems-new administrative, technological applications that are changing the ways in which records are created, stored, transported, and accessed. These systems require that registrars become familiar with new terminology, concepts, and technologies that, as discussed above, may not be fully understood even by those considered to be technical experts on campus.

Both system administrators and registrars have important areas of expertise. But it seems clear from this study that these individuals are not communicating with each other enough, if at all. Given differences in terminology, in understanding of data sensitivity and protection, in comprehension of the capabilities and operations of computer networks and systems, and in basic responsibilities, more interaction is required. Like individuals using the other's language as a "second" language, there are too many opportunities for missed communication and misunderstandings; only frequent and systematic exchanges and systems requiring checks and balances on these two disparate areas can rectify this situation.

Results: More Subtle Conclusions

This study highlights "obvious" findings as well as several that are more subtle. These deserve serious and immediate discussion on every college and university campus as well as subsequent action to better protect student records from potential privacy abuses; systems from misuse and abuse;

and institutions from increased liabilities. For each of these findings, the authors make a recommendation for action.

RESPONSIBILITY WITHOUT ADEQUATE INFORMATION

College and university staff members are handling data about which they do not have adequate information. In the networked, technology-dependent environments of college and university campuses, many responsible staff members are expected to manage resources without sufficient information or training to do so effectively, ethically, responsibly, and legally. If this was true of the system administrators in the LAMP study, consider how many others on college and university campuses—those not felt to be "most knowledgeable" about functions on the systems they manage—are operating without critical information about systems and sensitive data.

Recommendation: Colleges and universities must provide all system administrators with more adequate training in security on the systems they manage; in data protection; in fair information practice; and in relevant law, particularly in the rights afforded to students under FERPA. Provide instruction in FERPA confidentiality requirements and other relevant laws for all those with root access on systems.

DECISION RESPONSIBILITY WITHOUT FULL TECHNICAL UNDERSTANDING

College and university staff members are making decisions about applications without fully understanding the technological capabilities. When institutions require individuals who have expertise in one area/discipline to make decisions about technology applications they don't fully understand, both the individual and the institution find themselves in jeopardy. The individual may make decisions based on inadequate or incomplete information only to find that protections of data that were assumed are not in place, or that access rights they thought were limited in fact involve many other people. Registrars and others attempt to collaborate with their counterparts in the area of technology when they make decisions about enterprise applications. However, it often is the case that terminology differences and miscommunication, as well as time constraints, result in incomplete communication that can result in liabilities for the institution as a whole. Both registrars and system administrators were seen to be victims of this condition in the LAMP study.

Recommendation: Colleges and universities must provide registrars with more adequate training in security on the systems that handle and transport student data on campuses. Such training must include information about the logging and monitoring capabilities of new and emerging technologies. Registrars should define logs that have IP, domain name, user ID, date-time stamps, and any other identifiable information as education records. They should regularly define what education records exist and provide the name of the person responsible for maintaining the record. They should provide wider publicity and education for campus members concerning student record privacy rights, criteria for legitimate educational interest, and definitions of school official and what constitutes "need to know."

OVERLOAD AND FRUSTRATION OF PROFESSIONAL TECHNICAL STAFF

Many system administrators are under-trained and under-supported. Such conditions contribute to or even perpetuate failure. They can result in administrators taking actions that may prove faulty or inappropriate, particularly when time pressures are a factor. Unusual and inappropriate license may be taken and rules may be "bent" to accomplish tasks perceived as critical. System administrators need a clearer career path—one where additional training and more responsibilities are rewarded with fiscal compensation, at market value, and where other forms of compensation are readily available.

Recommendation: Colleges and universities must provide system administrators with a professional career path, one where more advanced training is tied directly to commensurate financial compensation and where advanced levels of responsibility are clearly coordinated with increased competency across multiple operating system platforms.

INADEQUATE PROTECTION FROM UNWITTING ACTS

Without adequate guidelines, policy, interpretations of law (where relevant), and standards of behavior, college and university staff are left unprotected from actions they may take in good faith but without adequate information. The absence of written policies about access to sensitive data, access and manipulation of log data, and even whether logs are subject to the protections of FERPA results in administrators of these data being vulnerable to mistakes or to committing acts that might compromise the privacy of users and increase institutional liability. Were it not for system administrators' own professional ethics and for established, albeit not written, practices, more difficulties surely would exist regarding access to log data. Responsibility and authority are not the same. When the presumption exists that authority is automatically tied to job description, it is easy for indi-

viduals to act beyond appropriate "authority" in particular situations. Rapid turnover can result in changed processes. The checks and balances that come with written policies and with defined levels of authority for obtaining additional privileges are absent. Staff members are vulnerable to making mistakes.

Recommendation: The Department of Education Family Policy Compliance Office must provide sample policies for institutions regarding access to student education record data in electronic environments. Sample scenarios also would be beneficial for all college and university community members. AACRAO already has been helpful in this regard. However, the material that has been written does not seem to be reaching the technical staff of colleges and universities. Registrars also have an obligation in this regard and must expand their efforts to define and educate individuals about privacy and their rights under FERPA. A warning should appear whenever a user logs on to a student information system, *e.g.*, "You are attempting to access legally protected information." Colleges and universities must establish effective processes for removing user access immediately when a user leaves the institution, and particularly for removing access privileges for those in positions that allow root access on machines. Colleges and universities must decrease the number of people who have access to log data and must set clear destruction schedules for data after they have been used. A complete process of audits should be put into place to identify each person who accesses sensitive or protected data.

DEFINING THE LEVEL AND LIMITS OF LOGGING

The activities system administrators must pursue for the sake of security and for operations and network management—logging and monitoring—must be considered at three levels in relation to privacy considerations.

Recommendation: The following three levels of logging activity are proposed for consideration as a method for matching data access with responsibilities and policy. Adoption of this or some similar system will assist administrators in establishing consistency and coordination.

LEVEL I

- Logging at Level I is for the purpose of network or operations management. Either data yielded cannot be associated with an individual user or functions are enabled in such a way as to effectively separate identifiable information from other output.
- Analysis and search rules are established to prevent individually identifiable information from being matched with output.
- This level of logging should be encouraged as often as possible, and logs should be analyzed regularly to maintain systems and operations.

LEVEL II

- Logging at Level II is also for the purpose of network and operations management as well as security. The data that are yielded may be associated with individual users through multiple steps. However, the data are separated into log outputs to facilitate analysis of specific functions but to provide checkpoints before data can be linked and related in such a way that education records are created.
- This level of logging should also be encouraged for the sake of system security and stability.
- The number of individuals with access to Level II log data should be restricted.
- Policies should be clearly written and disseminated regarding the linking of data and the conditions under which such can be done.
- Policies must designate the person who is the school official when data are combined as education records, and determinations need to be made relative to their access or release.
- Log data at Level II should be protected. Barriers-specified authorization processes-should be put into place in the event that data linking is necessary.
- Individuals handling data at this level must be trained in FERPA and must sign an agreement to abide by appropriate policy and relevant law.
- Archiving of Level II data is short in duration, with destruction following the designated period of use.

LEVEL III

- Logging at Level III is primarily for the purpose of security. Data yielded at this level include IP addresses, user IDs, account information, email addresses, date and time stamps, and other readily identifiable information. Individuals with access to this level log are very few in number and have high-level authorizations are documented in their position descriptions.
- Levels of authorization are canceled immediately when the individual changes jobs or leaves the agency.
- Individuals dealing with this level of data are highly trained in FERPA and data access procedures.
- The repercussions of violating data security levels are written, disseminated, and enforced.

- The annual notification to students clearly identifies this level of logs as education records, and students are given rights of review and correction of these records under FERPA.
- Archiving of combined student data is extremely short, and destruction of such data, after use, is prompt.
- "Legitimate educational interest" for access to these log data is defined in policies and operational guidelines.

INVESTIGATION AND PURSUIT WITHOUT SUFFICIENT COLLABORATIVE CONSULTATION

Because of the lack of policies and training regarding regulations, law, fair information practice, and data protection, college and university personnel are pursuing abusers of their systems and potential abusers of systems and data on their own. Many probably do not even realize that there are appropriate procedures for conducting such investigations, that evidence preservation is an important aspect of criminal investigations, or that simple techniques for investigating patterns of potential abuse may violate privacy rights and law. LAMP study respondents occasionally indicated that they would confer with university counsel or with staff members designated as user advocates or problem handlers before going further in gathering log information. However, in many cases, this was not considered necessary. The "bright line"—the line beyond which they should not go in gathering information on suspected system abusers—was difficult to see. In fact, the line may not even exist in the minds of many system administrators because they have received little help in drawing it. Absent the proper authority and guidance, investigation and discipline of students are particularly disturbing. Colleges and universities are environments designed for learning. Individual interventions designed to teach a student about his or her misbehavior and why it is a problem are more appropriate than surreptitious punishments determined without adequate consultation.

Recommendation: Many specifically trained individuals on college and university campuses can help establish a "bright line" for appropriate process. Attorneys, student affairs officers, technology incident/problem handlers, and campus safety officers may be of assistance in this area. Colleges and universities must establish and widely disseminate procedures that allow system administrators to take fast action to protect systems when they are in danger. The procedures must make clear with whom system administrators must consult and from whom they must obtain additional authority when investigations escalate to target specific individuals or involve potentially criminal actions.

A LOW ROAD TO FAIR INFORMATION PRACTICE

The LAMP study shows that knowledge of fair information practice is lacking. This is not the fault of any one group, and it certainly is not unique to the campuses that participated in the study. The well-established but currently little discussed principles of fair information practice seem to have been lost in the rush to implement information technologies on campuses. Like other institutions in our litigious culture, colleges and universities may be too quick to depend on attorneys to make determinations about policy governing the handling of personal information. Like the fox asked to guard the chicken house, attorneys often find themselves in a difficult position—one that too often results in an overly conservative stance. As protectors of the institution, they may recommend policies that over-protect the institution and under-protect individuals.

It seems clear that the principles of fair information are not being met. It is highly unlikely that notification-the requirement that an individual be informed when data about him or her is collected-is happening adequately; after all, there are places on campuses where log data are being collected without the understanding that such data constitute education records. Minimization—the principle according to which the minimum amount of data is collected to complete a required task—also is not happening; because of time constraints and lack of personnel, logs are being collected and stored without being analyzed. Secondary use—the principle that restricts data use to the purpose for which it was initially collected—is in danger of being compromised because of the absence of policies on access. Nondisclosure and consent-the principle that prevents information from reaching third partiers unless permitted by legal exception or consent cannot be effective until it is clear to all data handlers that log data constitute education records and thus are protected. Need to know—the principle that restricts access to data by individuals without clearly approved institutional responsibilities—is likely to be ineffective absent clear definitions of who is authorized to access such data. The lack of policy in this area creates liability for the institution and its personnel. Data accuracy, inspection, and review—a principle that encourages individuals about whom the data refer to enhance accuracy through inspection—can only occur if notification is fully and widely made. Finally, information security, integrity, and accountability—the principle that places on the data holder the responsibility to ensure that data are not altered without authorization—is the very principle that system administrators are trying to fulfill in their security and operations maintenance responsibilities.

However, without tighter controls than mere ties to job description and root access, this cannot be achieved.

Recommendation: Colleges and universities must have the courage to take the high road to fair information practice and the handling of personal information. They must put in place institutional policy that defines an individual's responsibility when releasing personal information to the institution and the institution's responsibility in collecting, transporting, using, and storing such information. Education about the principles of fair information practice must be widespread and provided annually. Individuals must be empowered to manage their personal information responsibly through information and continual choice. Institutions must fully disclose the personal information they are collecting and the purposes for which such information is being used, assured that fully informed and involved students will assist them in reaching the goals of valid and reliable information use. Penalties for abusing or misusing personal information must be enforced.

THE SURVEILLANCE CREEP OF TECHNOLOGY

Perhaps the most subtle and troubling of the findings of the LAMP study pertains to the growing potential of technology being used for surveillance purposes. Two results from the LAMP study pertain specifically to this discussion: (1) the trivial effort required for an individual who has root access on a computer he or she is managing to get from log data to the identity of an individual (typically one step); and (2) the percentage of individuals who have used logging applications to collect information about individual behaviors. We have already commented on the need for system administrators to investigate and stop potential abuses or attacks on their systems. Likewise, we have commented on the importance of layers of authority as checks and balances against unwitting acts or abuses of power against individuals.

Concern about surveillance creep—the tendency to increase the potential and range of surveillance capabilities—is justified and extends beyond the activities of system administrators on individual campuses. The use of surveillance technologies and the tendency to increase the power, scope, and interrelatedness of such technologies has significance for society in general. It must be consciously examined because it will affect the culture of the future.

James Rule et al. write, "Much organizational interest in the details of people's private lives relates to the effort to curtail one or another form of deviant behavior. Thus new forms of surveillance are especially likely where they promise to enable organizations to root out some troublesome form of misbehavior."[18]

The security concerns regarding computer systems and networks on campuses constitute fertile ground for the development of increased surveillance technologies. Another driving force is the desire to extend the technologies to their outer limits. This is inevitable, as colleges and universities are inherently teaching, learning, and experimental environments where such exploration is valued and supported.

Rule et al. identify three reasons for curtailing the unlimited growth of surveillance technologies: (1) They can result in excessive concentrations of social power in central organizations; (2) they can disenfranchise and disempower the population to which they are applied; and (3) they can render social relations excessively "unforgiving" by preserving data on people's past misdeeds which might better be forgotten.

Recommendation: Colleges and universities must openly and actively discuss the potential of surveillance technologies on campuses and the conditions that lead to their proliferation and growth, and they must make conscious community-based decisions about when and if such technologies will be curtailed. They have an obligation to lead the discussion of these technologies, for the sake of other organizations, for the public in general, for the cultures and futures now being created.

If colleges and universities are to be the open communities about which Ernest Boyer speaks—"a place where freedom of expression is uncompromisingly protected and where civility is powerfully affirmed"[19]—then individuals must understand and have a voice in the handling of their personal data. Moreover, if colleges and universities are to be the "just communities" about which Boyer speaks-places where the sacredness of each person is honored and where diversity is aggressively pursued"[20]—then we cannot be satisfied with procedures that violate the rights of individuals either because of inadequate policies and guidelines or because of a shallow and short-sighted notion of security.

18 Rule, J., D. McAdam, L. Stearns, and D. Uglow. 1980. The Politics of Privacy: Planning for Personal Data Systems as Powerful Technologies, New York: Elsevier.
19 Boyer, E. L. 1990. Campus Life: In Search of Community. The Carnegie Foundation for the Advancement of Teaching, Princeton, NJ.
20 Ibid.

Appendix A

Participating Schools

Sixteen colleges and universities participated in the study. (Fourteen of the college and universities' registrars participated.)

West Coast Participants

Santa Clara University
Stanford University
University of California at Berkeley

Midwest Participants

Northwestern University
University of Chicago
University of Illinois-Chicago

East Coast Participants

Binghamton University-
State University of New York (SUNY)
Cornell University
Syracuse University

Southern States Participants

Emory University
Georgia State University
Georgia Technical University

Additional Midwest Participants (Michigan)

Kalamazoo College
Michigan State University
Oakland Community College
Wayne State University

Appendix B
System Administrator Data Collection Instrument

Interviewer:
Date of interview:
Name of interviewee:
E-mail address:
Phone Number of interviewee:
Contact address:

Part A: Basic information about you:

- What is your job title?
  Briefly describe your responsibilities:
- How much professional experience do you have?
  ❑ 1 year ❑ 2–5 years ❑ 5–10 years ❑ 10 years
- How long have you been at your current job?
  ❑ 1 year ❑ 2–5 years ❑ 5–10 years ❑ 10 years
- What is your highest education level?
  ❑ H. S. ❑ A.A./A.S. ❑ B.S. ❑ M.S. ❑ Ph.D.
- In What field?
- Have you taken courses in systems security on the systems you administer?
  ❑ YES ❑ NO
- Do you have any technical certifications?
  ❑ NT ❑ Netware ❑ UNIX ❑ CISCO ❑ Other:
- Have you taken courses in fair information practice, data access, and data protection?
  ❑ YES ❑ NO
- Have you taken courses on compliance with FERPA (Family Educational Rights and Privacy Act) in handling student records?
  ❑ YES ❑ NO
- What type of systems are run at your university?
  ❑ NT ❑ LINUX ❑ NOVELL (Netware) ❑ UNIX
  ❑ HP ❑ AIX ❑ AS4OO ❑ oS390 ❑ OTHER:

Part B: Interviewee

- System
  ❑ NT ❑ LINUX ❑ NOVELL (Netware) ❑ UNIX
  ❑ Solaris ❑ HP ❑ AIX ❑ AS400 ❑ oS390
  ❑ OTHER:
- What is the primary function of this system?
- Are these services provided for a:
  Department? School? University system?
  State backbone?
- On a scale of 1 (not sensitive) to 4 (extremely sensitive), how sensitive do you think the data on this system are?
  ❑ 1 ❑ 2 ❑ 3 ❑ 4
- On a scale of 1 (no protection) to 4 (extensive protection), how much protection should such information have?
  ❑ 1 ❑ 2 ❑ 3 ❑ 4
- Do you run logs on this machine?
  ❑ YES ❑ NO
- For what purpose do you log?
- Is this default logging?
  ❑ YES ❑ NO Are some default functions turned off? If so, which ones?
- Do you also enable some functions?
  ❑ YES ❑ NO What specifically? Who or what is the target? What is the purpose?
- Have you created specific customized logging (scripts)?
  ❑ YES ❑ NO What specifically? Why?
- What type of data does your logging yield?
  ❑ IP addresses ❑ Domain names
  ❑ Unique names/USER IDs ❑ Time stamps
  ❑ Student information Personnel data
  ❑ Mail ❑ OTHER:

- What and how many steps would you need to go through to get from these data to a person's name?

- How often do you analyze the logs? Specifically:
  ❑ Daily ❑ Weekly ❑ Monthly ❑ OTHER:

- What patterns are you analyzing for?

- Do you target specific activities/patterns for individuals and collect logs when those characteristics occur?
  ❑ YES ❑ NO

- Do you archive the log information that you collect?
  ❑ YES ❑ NO Why?

- If so, for how long? Specifically how long:
  ❑ 3 days ❑ 4–7 days ❑ 1 week–2 months ❑ 2months

- Is authorization required to access these data (re: logs)?
  ❑ YES ❑ NO

What criterions govern access?

- How many people can access these data?
  ❑ 5 people ❑ 5–10 people ❑ 11–25 people ❑ 25 people

- Do policies/standards exist that govern access?
  ❑ YES ❑ NO What?

- On a scale of 1 (not serious) to 4 (extremely serious), how serious would it be for someone to access these data without authorization?
  ❑ 1 ❑ 2 ❑ 3 ❑ 4

- On a scale of 1 (not difficult) to 4 (extremely difficult), how difficult is it to get to data without access?
  ❑ 1 ❑ 2 ❑ 3 ❑ 4

- What barriers exist that prevent you from such access?
  ❑ Knowledge ❑ Time ❑ Skill ❑ Authorization
  ❑ Equipment Policy ❑ Law Personal ethics

- What more logging would you like to do in the future?

- What barriers exist that prevent you from doing so?
  ❑ Knowledge ❑ Time ❑ Skill ❑ Authorization
  ❑ Equipment Policy ❑ Law Personal ethics

Appendix C
Sample System Administrator Job Description

Senior System Administrator III

DUTIES:

- Administer, maintain, and operate large departmental computing environments for the school and national and international customers in multi-platform environments. These environments will consist of Windows NT-4 server and workstation, Windows NT-5, Novell, Solaris and SunOS, Linux, and McIntosh. Interoperability among all of these platforms is required.
- Resolve complex system hardware, software, and networking failures. Train support staff.
- Take a lead role in designing, negotiating, developing, and building operational infrastructure for future networked architectures for the school, the university, and for consortia of related organizations.
- Automate routine tasks. This will include extensive involvement deploying combinations of externally developed Web-based collaboration products.
- Negotiate, implement, and lead the operation of security systems and processes.
- Arrange performance, usage, and security monitoring; optimize systems based on these data.
- Provide administrative supervision for one or more staff.
- Ensure compliance with affirmative action and computer security and privacy programs.

MINIMUM QUALIFICATIONS:

- M.S. in Engineering or Computer Science or equivalent;
- Ten or more years of progressively responsible and hands-on experience administering Unix, Windows NT, and similar systems;
- Demonstrated team leadership experience;
- Experience administering and/or developing advanced HTTP services;
- Excellent, demonstrated verbal and written communication skills.

DESIRED QUALIFICATIONS:

- Ten years' experience administering Unix systems, including team leadership in this role.
- Solaris, Sun/OS, and Linux administration experience.
- Two or more years' experience administering Novell, including integration of Novell with other systems, including Unix.
- Multimedia Web development and Web administration experience.
- Working knowledge of the university's infrastructure.
- Experience being responsible for system support for a public computer laboratory.
- Successful experience writing grant proposals.
- Experience coordinating national and international networking efforts.
- Experience working and negotiating with vendors to produce and implement open solutions

Appendix D
Glossary of Terms

Archiving
Focus Group Definition: Archiving is the action of recording and storing copies of logged data for the purpose of freeing up disk space and for analysis of incident investigation or for a record of the system's operational history.

Backup
Focus Group Definition: The term "backup" refers to the practice of recording the system state and current data in case of potential system failure. In the event of a system procedural error, a system "backup" on media such as magnetic tape can be used to restore the system to its last known functional state.

Default Logging: A data collection function provided directly by vendors at the time the particular system is configured by the manufacturer. Focus Group Definition: Default logging is considered to be the recording configurations that are predefined and set by the manufacturer of the software tool. This is the state in which the system arrives as it was originally packaged by the designer and manufacturer, with no adjustments made by the system administrator to suit his or her particular operation. Although most systems arrive with default logging set, some manufacturers distribute their systems with the logging function turned off, which makes it questionable as to whether it can be considered the systems default setting. Use of the phrase default logging drew a mixed response from our participants simply because the answer varies across systems.

Directory Information: Student information maintained by the education institution and which may include such information as the student's name, address, telephone number, date and place of birth, major fields of study, participation in officially recognized activities and sports, weight and height (if an athletic team member), photograph, dates of attendance, degrees and awards received, most recent education institution attended, and other information as defined by the institution which generally would not be considered harmful to the student or an invasion of privacy if disclosed.

Disciplinary Action or Preceding: Investigation, adjudication, or imposition of sanctions by an education institution with respect to an infraction or violation of the internal rules of conduct applicable to students of the agency or institution.

Disclosure: Permitting access to or the release, transfer, or other communication of education records of the student or personally identifiable information contained therein to any party orally, in writing, by electronic means, or by any other means.

Domain Name System: Computer naming system developed by the National Science Foundation for denoting the names given to each computer registered to a particular system, corporation, institution, etc.

Education Records: Records directly related to a student and maintained by the institution or by a party acting for the institution. The term education records does not include the following:

- Records of instructional, supervisory, administrative, and certain educational personnel which are in the sole possession of the maker thereof and which are not accessible or revealed to any other individual except a substitute who performs on a temporary basis (as defined in the institutional personnel policy) the duties of the individual who made the records.
- Records maintained by a law enforcement unit of the educational agency or institution that were created by that law enforcement unit for the purpose of law enforcement.
- Records relating to individuals who are employed by the institution and which are made and maintained in the normal course of business, relate exclusively to individuals in their capacity as employees, and are not available for any other purpose. (Records of individuals in attendance at an institution who are employed as a result of their status as students are educational records, *e.g.*, work-study.)
- Records relating to a student which are (1) created or maintained by a physician, psychiatrist, psychologist, or other recognized professional or paraprofessional acting in his or her professional capacity or assisting in a paraprofes-

sional capacity; (2) used solely in connection with the provision of treatment to the student; and (3) not disclosed to anyone other than individuals providing such treatment, so long as the records can be personally reviewed by a physician or other appropriate professional of the student's choice. (The institution can determine appropriateness.) Treatment in this context does not include remedial educational activities or activities which are part of the program of instruction at the institution.

- Records of an institution that contain only information relating to a person after he is no longer a student at the institution (*e.g.*, information on the accomplishments of alumni).

Enabled Logging: Data collection function that is provided with the system but which has to be turned on or otherwise reconfigured to provide the desired log data designated by the system administrator. Focus Group Definition: Enabled logging is the action of reconfiguring the logging system to change the amount or type of information yielded, such as receiving more logging data or filtering logs for specific information. This ability allows the system administrator to develop system logs with varying degrees of detail, depending on what has been selected or enabled

IP Address: Address assigned to every computer linked to the Internet that uniquely identifies it as that hardware unit while on the network

Law Enforcement Unit: Any individual or other component of an institution, including commissioned police officers and noncommissioned security guards, officially authorized by the institution to enforce any local, state, or federal law and to maintain the physical security and safety of the institution. (Although the unit may perform other non—law enforcement functions, it does not lose its status as a law enforcement unit.)

Legitimate Educational Interest: The demonstrated need to know by those officials of an institution who act in the student's educational interest, including faculty, administration, clerical, and professional employees, and other persons who manage student record information. (Although FERPA does not define "legitimate educational interest," it states that institutions must establish their own criteria, according to their procedures and requirements, for determining when their school officials have a legitimate educational interest in a student's education records. This is a recommended definition. Sample institutional statements on legitimate educational interest are included in FERPA Appendix II.)

Log: Collection of information recorded for study and analysis, in this case referring to computer network events.

Logging: Process of systematically or automatically collecting information and recording it to a detailed document for later study and analysis.

Monitoring: Process of systematically or automatically watching or responding to patterns that occur. In computer monitoring, this generally means that designated patterns seen on the logs are responded to either automatically or personally as indications of operational problems or as problems caused by users.

Personally Identifiable: Data or information which include (1) the name of the student, the student's parent, or other family members; (2) the student's address; (3) a personal identifier such as a social security or student number; or (4) a list of personal characteristics or other information that would make the student's identity easily traceable.

Record: Any information or data recorded in any medium (*e.g.*, handwriting, print, tapes, film, microfilm, microfiche, or any form of electronic data storage.)

School Officials: Members of an institution who act in the student's educational interest within the limitations of their need to know. These may include faculty, administration, clerical, and professional employees and other persons who manage student education record information. (Although FERPA does not define "school officials," it states that institutions must establish their own criteria, according to their own procedures and requirements, for determining who they are. This is a recommended definition.)

Scripted Logging: Data collection function that is not provided with the system automatically but which is designed and installed by the system administrator by writing code to direct the system in the collection, sorting, or transporting of particular data according to the instruction set provided. Focus Group Definition: The term customized/scripted logging also received mixed responses from the research participants. Depending on the type of system on which the logging occurs, the answer may vary. According to the project technical focus group, customized logging is considered the pre-processing of log data or the generation of non-standard software to produce custom messages or warnings.

Student: Any individual for whom an education institution maintains education records. The term does not include an individual who has not been in attendance at the institution. An individual who is or who has been enrolled in one component unit of an institution and who applies for admission to a second unit has no right to inspect the records accumulated by the second unit until he or she is enrolled therein.

Appendix E:
Percentage of Respondents Who Yield Data Types

Figure 1: Data Yield—IP

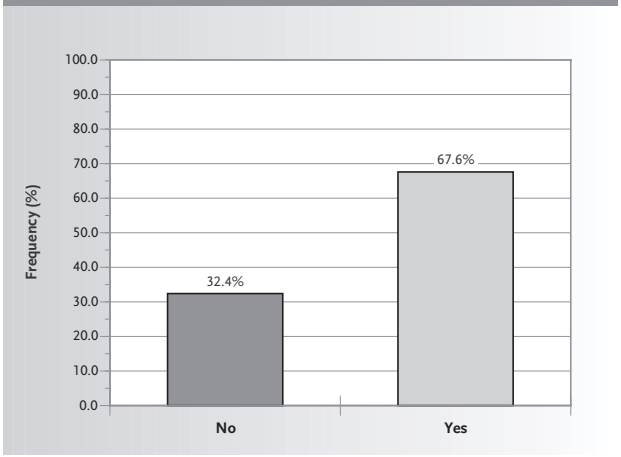


Figure 2: Data Yield—Date, Time

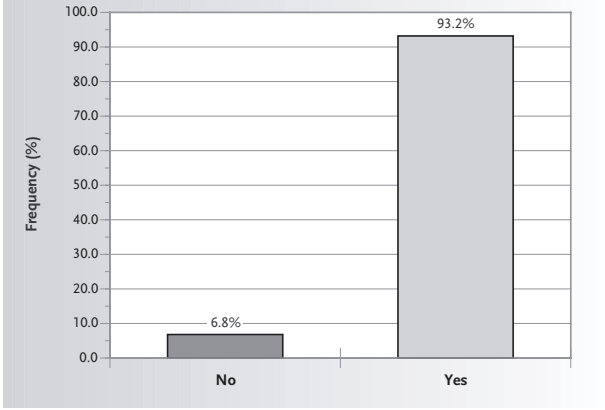


Figure 3: Data Yield—Domain Name

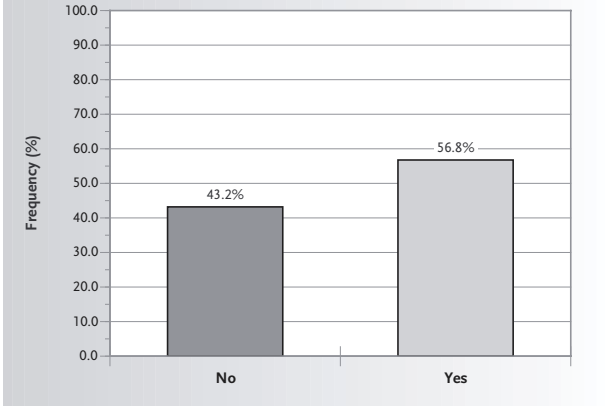


Figure 4: Data Yield—Unique Name

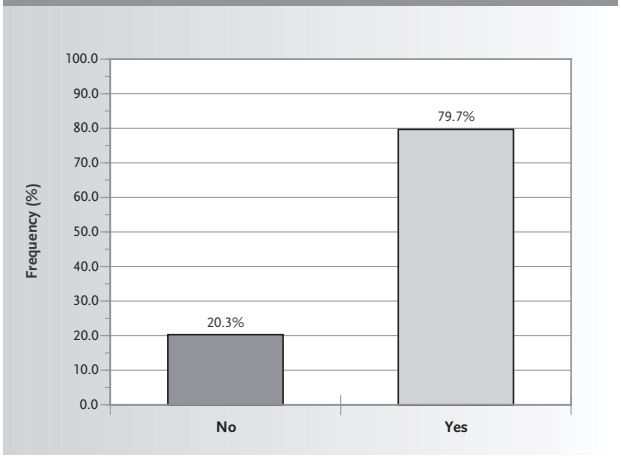
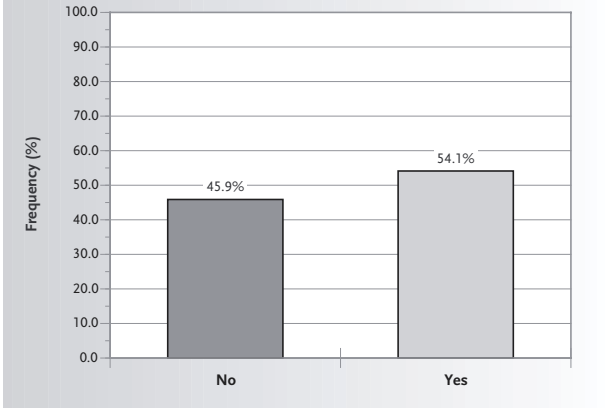


Figure 5: Data Yield—Other



Appendix F
Registrar Scenarios and Data Collection Instrument

Scenario 1: Let Me Know

John, a system administrator in a college math department, collects information on the operation of the department's computers and networks. He needs to watch the flow of traffic over his systems to responsibly manage the electronic resources. Students, faculty, and staff heavily use the computers. John's log information includes the number of people who sign on, the number of people using mail at any given time, the number of packets of information sent across the networks, and the static IP addresses of the machines from which, and to which, those packets are sent. Generally, John does not collect information about the kinds of services individual users seek, where they go on the Web, how much time they spend e-mailing or downloading files, or specifically what files they look at. However, today is different. The Dean wants to know how three student staff members are using their time. He has demanded that information be collected about them and their use patterns from their machines and turned over to him for analysis. John secretly turns on the logging function and collects information from the machines used by these students. The information includes time of access, the specific services to which the student staff members go, what content they are viewing, how long they were at particular sites, what e-mail and to whom e-mail was sent, and when they logged off their machines.

QUESTIONS:

❑Yes ❑No a] Does any of the information referred to in this scenario constitute "an educational record" under FERPA? If "Yes," which ones and why? If "No," why not?

❑Yes ❑No b] Does collection of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No c] Does the sharing of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No d] Is this an appropriate "educational use" of student information under FERPA? Why or why not? Please explain:

❑Yes ❑No e] Does access to this information qualify under "legitimate educational interest?" Why or why not? Please explain:

Additional comments:

Scenario 2: 24x7 Information

Tudlo College Is very electronically advanced. The staff members are particularly proud of the fact that all of their computer resources have been networked together using state-of-the-art technology. If a staff member needs Information regarding a particular student they can obtain It quickly and efficiently. The Information they store is extensive. Not only do they have student grades, addresses, course selection lists, financial aid and other important information online, but the college has now begun to collect Information about student transactions as well. Through their extensive networks, the use of card swipe machines and key card mechanisms across campus, information regarding student transactions is collected twenty-four hours each day and seven days per week. This allows the staff to know when a student leaves his residence hall, which building he or she enters and when, if they enter a particular room in that building and even when they leave. It provides information about if and when the student makes a purchase, eats a meal, reserves a book at the library, even where they park their car and how long they stay in a particular parking space. Each time the student swipes their ID card, data are collected and stored in the college's databases. In the event of an emergency, the college will be able to locate a student rapidly and with great certainty. They feel certain that student safety will be enhanced by this new capability.

QUESTIONS:

❑Yes ❑No a] Does any of the information referred to in this scenario constitute "an educational record" under FERPA? If "Yes," which ones and why? If "No," why not?

❑Yes ❑No b] Does collection of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No c] Does the sharing of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No d] Is this an appropriate "educational use" of student information under FERPA? Why or why not? Please explain:

❑Yes ❑No e] Does access to this information qualify under "legitimate educational interest?" Why or why not? Please explain:

Additional comments:

Scenario 3: Complaint Follow-up

Fred is a system administrator for a very large college In a mid-western university. He receives a complaint from one of the college's sophomores. The student, Sarah, describes a series of email messages that she has received from a person whose name she does not know. Though the name on the email is not familiar, the pattern of communication is. Sarah feels certain that she knows who Is sending her these repeated messages. She names Kenny, a senior in the college, as a person who Is stalking her. She wants Fred to take action as soon as possible. Sarah reports that even when told to stop sending her email, the sender does not stop and in fact increases the demands in his messages. He tells her what time her first class is on campus, which building she goes to in the morning, and where and when she eats meals. Sarah Is very frightened and is having difficulty studying because of her nervousness. From system logs, Fred can identify the network, the machine, and even the account from which the email message was sent. He cannot be sure however that it was Kenny who was using this account at the time the mail messages was sent. He is certain however that the account is in Kenny's name. Sarah wants action. Fred has some information that points to Kenny. He decides to follow-up by intensifying the collection of information on Kenny's machines. He writes a script that will alert him each time this account is in use, and will provide information on the exact services that Kenny uses, displaying the email and file content for any transaction from this account Additionally, he establishes a system with the residence hall advisors to tell him when Kenny is in his room, and therefore using the IP address to which his machine is assigned. He will soon have a large quantity of information about Kenny's electronic activities. His plan is to watch Kenny's online behavior, analyzing log records for at least two months.

QUESTIONS:

❑Yes ❑No a] Does any of the information referred to in this scenario constitute "an educational record" under FERPA? If "Yes," which ones and why? If "No," why not?

❑Yes ❑No b] Does collection of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No c] Does the sharing of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No d] Is this an appropriate "educational use" of student information under FERPA? Why or why not? Please explain:

❑Yes ❑No e] Does access to this information qualify under "legitimate educational interest?" Why or why not? Please explain:

Additional comments:

Scenario 4: Campus Safety Needs It

As the University becomes more electronically collected, Sgt. Denver, an officer with the University's Department of Campus Safety, is delighted with the potential information that will be available. Students are now asked to have their pictures digitized for inclusion on ID cards. They use keycard access for entry and departure from buildings, their unique names and machine IP addresses are stored in readily accessible university databases, and their course and meal-time schedules are made available to the campus safety staff as well. Sgt. Denver particularly likes the availability of the pictures online. This is information that the campus needs. Through the use of cameras, these images can be matched with individuals entering campus buildings, or those stopped for suspicious behaviors on campus and their identities can be verified. If they do not belong on the campus, they can be readily identified as non-students. If they are students but in locations without authorizations they can be notified or reprimanded. Pictures can even be used to show individuals who have witnessed a crime in order to identify suspects and provide "virtual lineups" to assist the officers in their work. The picture information can be connected to applications for employment or to resumes being sent to graduate or professional schools. College departments are already using them to assist professors in identifying the

students in their large classes. Keeping this information online, allows for its easy access and rapid transit when needed by different organizations on campus and officers in their different locations. It can be kept available indefinitely.

QUESTIONS:

❑Yes ❑No a] Does any of the information referred to in this scenario constitute "an educational record" under FERPA? If "Yes," which ones and why? If "No," why not?

❑Yes ❑No b] Does collection of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No c] Does the sharing of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No d] Is this an appropriate "educational use" of student information under FERPA? Why or why not? Please explain:

❑Yes ❑No e] Does access to this information qualify under "legitimate educational interest?" Why or why not? Please explain:

Additional comments:

Scenario 5: Better Watch This One

Sydney is in charge of all network systems at a small liberal arts college. He also manages all of the central services and servers for the college. It is very difficult to keep everything up-to-date, working together, and all systems coordinated. Sydney is not happy with the number of interruptions caused by unknown hackers who seem to explore systems looking for unprotected files and applications. He has configured his machines to log all machine activities, the amount of traffic flowing over the network, the number of machine errors that occur, and to notify him and other system administrators when machine problems occur. But this information does not tell them much about individual computing behaviors. Sydney knows that there are several students on the campus who have extraordinary computer skills and pride themselves in being able to do remarkable electronic feats on the networks. While they have never been identified for causing a computer problem or disciplined for inappropriate campus behavior, he decides that it is better to periodically watch these young men than remain ignorant of their activities. He selects one in particular who he thinks would be likely to engage in hacker activity and intensifies the information gathering logs connected with the student's ID, unique name, account use, the IP address of his machine, and all connections made from his residence hall. By periodically gathering and analyzing this extensive information about the student's local and web activity throughout the year, Sydney can be assured that this student is not participating in hacker activity, or if he is, can be quick to suspend his account or report him to student affairs for discipline.

QUESTIONS:

❑Yes ❑No a] Does any of the information referred to in this scenario constitute "an educational record" under FERPA? If "Yes," which ones and why? If "No," why not?

❑Yes ❑No b] Does collection of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No c] Does the sharing of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No d] Is this an appropriate "educational use" of student information under FERPA? Why or why not? Please explain:

❑Yes ❑No e] Does access to this information qualify under "legitimate educational interest?" Why or why not? Please explain:

Additional comments:

Scenario 6: Print Server Logs

The administrators at one central university are growing alarmed by the cost of providing computing resources to the campus community. They are struggling to find ways to pass the costs of particular services back to the students, and in some way to "throttle" the overuse of resources that they suspect may be happening. To do this they must have access to name-related use patterns on specified services. Systems administrators are instructed to initiate logging mechanisms on all print-servers in the public computing sites. The data logs provide information about the name of the student printing material, their unique name and account number, the number of sheets of paper printed, the time and date of

the printing, whether the printing was graphics or text type, and the machine from which the printing request was sent. The logs are analyzed and a bill is sent to the student for printing services. Logs are kept for two to four years for purposes of accounting.

QUESTIONS:

❑Yes ❑No a] Does any of the information referred to in this scenario constitute "an educational record" under FERPA? If "Yes," which ones and why? If "No," why not?

❑Yes ❑No b] Does collection of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No c] Does the sharing of this information constitute a violation of FERPA? Why or why not? Please explain:

❑Yes ❑No d] Is this an appropriate "educational use" of student information under FERPA? Why or why not? Please explain:

❑Yes ❑No e] Does access to this information qualify under "legitimate educational interest?" Why or why not? Please explain:

Additional comments:

Bibliography

Boyer, Ernest. 1990. *Campus Life: In Search of Community*. The Carnegie Foundation for the Advancement of Teaching, Princeton, NJ.

Dijker, B. (Ed.), 1996. A Guide to Developing Computing Policy Documents. *Short Topics in System Administration* #2. USENIX Association for SAGE, the System Administrators Guild. Berkeley, CA.

Family Educational Rights and Privacy Act, Final Rule, 34 CFR Part 99 (OS) http://www.ed.gpv/legislation/FedRegister/finrule/2000-3/070600a.html

Guidelines for Postsecondary Institutions for Implementation of the Family Educational Rights and Privacy Act of 1974 *as Amended*. 1995. AACRAO-American Association of Collegiate Registrars and Admissions Officers, Washington, DC.

Oppenheimer, D., D. Wagner, and M. Crabb. 1997. System Security: A Management Perspective. *Short Topics in System Administration* #3. USENIX Association for SAGE, The System Administrators Guild. Berkeley, CA.

Rainsberger, R., E. Baker, D. Hicks, B. Myers, J. Noe & F. Weese. 2001. *The AACRAO 2001 FERPA Guide*. AACRAO-American Association of Collegiate Registrars and Admissions Officers. Washington, DC.

Rezmierski, V., A. Carroll, and J. Hine. 1999. *Incident Cost Analysis and Modeling Project, I-CAMP II*. Final Report to the USENIX Association, Ann Arbor, MI.

Rezmierski, V., S. Deering, A. Fazio, and S. Ziobro. 1998. *Incident Cost Analysis and Modeling Project, I-CAMP I*. Final Report to the Committee on Institutional Cooperation Chief Information Officers Committee. Ann Arbor, MI.

Rezmierski, V., S. Ferencz, L. Alvarez, C. Goldsmith, K. Kimball, R. Morley, T. Pham, R. Smith, & S. Worona. 1997. *Privacy and the Handling of Student Information in the Electronic Networked Environments of Colleges and Universities*. CAUSE—The Association for Managing and Using Information Resources in Higher Education. Boulder, CO.

Rule, J. 1973. *Private Lives and Public Surveillance*. Allen Lane, London.

Rule, J., D. McAdam, L. Stearns & D. Uglow. 1980. *The Politics of Privacy: Planning for Personal Data Systems as Powerful Technologies*. Mentor Publishers, New York.

Schultz, E. Eugene. 2000. *Windows NT/2000 Network Security*. Macmillan Technical Publishing, USA.


270