ies NATIONAL CENTER FOR EDUCATION STATISTICS
Institute of Education Sciences

SLDS Technical Brief

Guidance for Statewide Longitudinal Data Systems (SLDS)

November 2010, Brief 1
NCES 2011-601

Basic Concepts and Definitions for Privacy and Confidentiality in Student Education Records

Contents

SLDS Technical Briefs are intended to provide "best practices" for consideration by states developing Statewide Longitudinal Data Systems.

For more information, contact:
Marilyn Seastrom
National Center for Education Statistics
(202) 502-7303
Marilyn.Seastrom@ed.gov

The National Center for Education Statistics (NCES) is launching a new series of Technical Briefs on various aspects of the protection of personally identifiable information in students' education records. The immediate demand for this work arose from increased federal mandatory reporting (20 U.S.C. § 6311) and the related expansion of record keeping under the Statewide Longitudinal Data Systems (SLDS) (20 U.S.C § 9607; Public Law 111-05 American Recovery and Reinvestment Act of 2009 (ARRA)). This increase in the amount of data published and stored must be balanced against the legal requirements under the Family Educational and Privacy Rights Act (FERPA) to protect personally identifiable information in student education records (20 U.S.C. § 1232g). (Education records include those records that are directly related to a student and are maintained by an educational agency or institution or by a party acting for the agency or institution (34 CFR § 99.3).)

While driven by recent events, the principles and practices that are outlined in this series can be applied more generally to personally identifiable information about students. This series of Technical Briefs is intended to be useful for anyone responsible for the development, maintenance, protection, or use of student record data. This first brief discusses basic concepts and definitions that establish a common set of terms related to the protection of personally identifiable information, especially in education records.

Personally Identifiable Information

The definition of *personally identifiable information* is central to all discussions of privacy and confidentiality. The Office of Management and Budget (OMB) Guidance for the implementation of the Confidential Information Protection and Statistical Efficiency Act of 2002 and OMB Memorandum 07-16, *Safeguarding Against and Responding to the Breach of Personally Identifiable Information,* both state that "The term 'personally identifiable information' refers to information that can be used to distinguish or trace an individual's identity, such as their name, Social Security Number, biometric records, etc. alone, or when combined with other personal or identifying information which is linked or linkable to a specific individual, such as date and place of birth, mother's maiden name, etc."

The National Institute of Standards and Technology (NIST) definition in the 2010 publication *Guide to Protecting the Confidentiality of Personally Identifiable Information* (NIST Special Publication 800-122, p. E-1) parallels the OMB definition. Although there is some variation in the wording of the definition across different applications of the term, the OMB definition is the basis for the definition of personally identifiable information across the federal government.

The Family Educational Rights and Privacy Act (FERPA) 2008 regulations (34 CFR § 99) define personally identifiable information for education data and student education records.

Personally identifiable information, as defined in FERPA, includes, but is not limited to:

1. The student's name;
2. The name of the student's parent or other family members;
3. The address of the student or student's family;
4. A personal identifier, such as the student's Social Security Number, student number, or biometric record;
5. Other indirect identifiers, such as the student's date of birth, place of birth, and mother's maiden name;
6. Other information that, alone or in combination, is linked or linkable to a specific student that would allow a reasonable person in the school community, who does not have personal knowledge of the relevant circumstances, to identify the student with reasonable certainty;
7. Information requested by a person who the educational agency or institution reasonably believes knows the identity of the student to whom the education record relates. (34 CFR § 99.3)

Discussions of personally identifiable information frequently use the concepts of *identifiable form* and *direct and indirect identifiers*. The first of these terms was codified in law in the E-Government Act of 2002 (Public Law 107-347). Section 208(d) of that Act states that "In this section, the term 'identifiable form' means any representation of information that permits the identity of an individual to whom the information applies to be reasonably inferred by either direct or indirect means" (44 U.S.C. § 3501, note).

The FERPA definition of personally identifiable information calls out specific direct identifiers, such as name, biometric record, Social Security Number, and student number. The FERPA regulations define a biometric record as including measurable biological or behavioral characteristics, such as fingerprints, retina and iris patterns, voiceprints, DNA sequence, facial characteristics, and handwriting (see 34 CFR § 99.3 for full definition).

The FERPA definition refers to "other indirect identifiers such as the student's date of birth, place of birth, and mother's maiden name" and to "other information that, alone or in combination, is linked or linkable to a specific student…" The FERPA definition also includes targeted requests—that is requests where the person requesting the information is trying to get information on a specific student. For example, if there was a rumor published in the local paper that a public official was disciplined for cheating during his senior year in high school, a request to the high school for the disciplinary records of students who were caught cheating during the year the public official was a senior would be considered a targeted request.

OMB Memorandum M-03-22 *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002* and OMB Memorandum M-07-16 *Safeguarding Against and Responding to the Breach of Personally Identifiable Information* provide additional examples of direct and indirect identifiers. Direct identifiers include information that relates specifically to an individual such as the individual's residence, including for example, name, address, Social Security Number or other identifying number or code, telephone number, e-mail address, or biometric record. *Indirect identifiers* include information that can be combined with other information to identify specific individuals, including, for example, a combination of gender, birth date, geographic indicator, and other descriptors.

The 2010 NIST guide extends the list of examples of indirect identifiers to include place of birth, race, religion, weight, activities, employment information, medical information, education information, and financial information (NIST 2010 Special Publication 800-122, p. 2-2).

FERPA allows the public release of some personally identifiable student information as school directory information (20 U.S.C. § 1232g

(b)(1)), where *directory information* is defined in the 2008 FERPA regulations as "information contained in an education record of a student that would not generally be considered harmful or an invasion of privacy if disclosed" (34 CFR § 99.3). The FERPA regulations also specify that directory information may not include a student's Social Security Number or an identification number that is used to access the student's education record. The FERPA regulations require that educational agencies or institutions provide public notice to parents of students or eligible students of the types of personally identifiable information that are designated as directory information (34 CFR § 99.37). The parent or the eligible student must be given the right to refuse to have any or all of the student's information released as directory information. (An "eligible student" is a student who has reached 18 years of age or is attending a postsecondary institution (34 CFR § 99.3).)

The 2008 FERPA regulations state that "Directory information includes, but is not limited to, the student's name; address; telephone listing; electronic mail address; photograph; date and place of birth; major field of study; grade level; enrollment status (e.g., undergraduate or graduate, full-time or part-time); dates of attendance; participation in officially recognized activities and sports; weight and height of members of athletic teams; degrees, honors and awards received; and the most recent educational agency or institution attended" (34 CFR § 99.3).

In the Notice of Proposed Rule Making (NPRM) for the 2008 FERPA regulations, the U.S. Department of Education recognized that the risk of identifying a student in aggregate data is cumulative and related to previous releases of data from student education records in both directory information and aggregate reports that are assumed to protect personally identifiable information, as well as to data from external sources. Furthermore, in acknowledging that these risks have increased as a result of new technologies and methods that emerged since FERPA was enacted in 1974, the Department advised "that parties should minimize information released in directories to the extent possible" (73 Fed. Reg. 15574-602, March 24, 2008).

Privacy and Confidentiality

The terms privacy and confidentiality are often invoked in discussions about rights and responsibilities when it comes to student records; in fact, they are often used interchangeably even though they have distinct meanings. So exactly what does each of these terms mean?

Privacy Defined

The concept of *privacy* relates to individual autonomy and each person's control over their own information (Report of the National Academy of Science 1993 Panel Report *Private Lives and Public Policies,* p. 3). This includes each person's right to decide when and whether to share personal information, how much information to share, and the circumstances under which that information can be shared (Report of the National Academy of Science 1993 Panel Report *Private Lives and Public Policies,* p. 22).

The 2009 National Academy of Sciences Report from the Committee on National Statistics and the Center for Education Workshop, *Protecting Student Privacy and Facilitating Education Research,* defined privacy as "…an individual's control over who has access to information about him or her. The concept of privacy is relevant to what personal information becomes data" (Summary of the Committee on National Statistics' 2009 Workshop on *Protecting Student Records and Facilitating Education Research,* p. 3).

In the context of student education records and FERPA, privacy pertains to the rights of parents and eligible students to inspect and review the students' education records, to seek to amend education records, to consent to the release of personally identifiable information from education records for any disclosures that are not authorized in law, and to refuse to have personally identifiable information that is designated as directory information publicly released (20 U.S.C. § 1232g, 34 CFR §§ 99.7, 99.37).

Confidentiality Defined

Confidentiality relates to the management of another individual's personally identifiable information. In a 2009 National Academy of Sciences, Institute of Medicine report, confidentiality is defined as referring to the obligations of those who receive personal information about an individual to respect the individual's privacy by safeguarding the information (Committee on Health Research and the Privacy of Health Information: The HIPAA Privacy Rule, 2009, *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research,* pp. 17–18).

These concepts are echoed in the 2009 National Academies of Sciences workshop report on protecting student records in which confidentiality is defined as "protection against the release of personal information. An important distinction is that privacy pertains to individuals; confidentiality to their information" (Report of the National Academy of Science 2009 Workshop *Protecting Student Privacy and Facilitating Education Research,* p. 4).

Legal and ethical responsibilities to protect against the release of personal information must be respected and enforced even if some of the same information is already in the public domain. The fact that some of the information is already in the public domain can make use or disclosure of other information more sensitive.

Disclosures of Confidential Information

These definitions introduce the concept of protecting personally identifiable information from release. This is also referred to as protecting personally identifiable information from *disclosure*. Under FERPA, "Disclosure means to permit access to or release, transfer, or other communication of personally identifiable information contained in education records by any means including oral, written, or electronic means, to any party except the party identified or the party that provided or created the record" (34 CFR § 99.3).

There are three types of disclosure—*authorized, unauthorized, and inadvertent*. FERPA *authorizes* or permits specific users and uses of personally identifiable information in student education records without the written consent of the parent or eligible student. These authorized disclosures include, but are not limited to, the following:

- other school officials, including teachers within the agency or institution who have legitimate educational interests;
- officials of another school, school system or postsecondary institution in which the student seeks to enroll;
- authorized representatives of the Comptroller General of the United States, Attorney General of the United States, the Secretary of the U.S. Department of Education, and state and local educational authorities;
- in connection with financial aid for which a student has applied or received;
- State and local officials or authorities to whom access is granted under state statute;
- organizations conducting studies for, or on behalf of, educational agencies or institutions for the purpose of developing, validating, or administering predictive tests, administering student aid programs, and improving instruction, subject to confidentiality and privacy conditions (including a written agreement);
- accrediting organizations for accrediting purposes;
- parents of a dependent student;
- information designated as directory information;
- a parent of a student who is under age 18 and not enrolled in postsecondary education;
- a student who has reached age 18 or enrolled in postsecondary education;
- in connection with a health or safety emergency (see 34 CFR § 99.31 for additional details and exceptions).

An *unauthorized* disclosure occurs when personally identifiable information from a student's education record is made available to a third party who does not have legal authority to access the information. An *inadvertent* disclosure occurs when information about an individual is unintentionally revealed through

information released to the public. This might happen, for example, through a security breach of the electronic system that is used to maintain and access the education records, as a result of a teacher or administrator leaving paper reports that include personally identifiable information in an unsecured location, or as a result of identifiable information about a student that can be derived from published summary statistics that were not fully protected.

The National Academy workshop report on protecting student privacy makes a further distinction between the confidential information in a student record that includes personal information and statistical reports derived from that information. The report cites as an example the fact that while a parent has the right to control information pertaining to the fact that his or her child is enrolled in a specific school, a summary statistic of the number of students enrolled in a school does not violate confidentiality and thus does not constitute a disclosure. In other words, it is not a disclosure or a violation of the confidentiality of the information in the data when personal information for a number of students is combined to produce a statistical report (Report of the National Academy of Science 2009 Workshop *Protecting Student Privacy and Facilitating Education Research,* p. 4).

While this is true in the case of a summary enrollment count, it is important to understand that even with statistical reports care must be taken to avoid *inadvertent* disclosures. Disclosures of this type are unintentional and occur when data in a student level file or aggregate data in tabulations allow the data user to identify a student, known as *identity disclosure,* or when data in a student level file or aggregate data in tabulations reveals sensitive information about a student, known as an *attribute disclosure.* For example, a statistical report of student assessment results for Hispanic third-graders in a specific school shows that there were students in this subgroup who scored in each of four different achievement levels. Knowing that these students were distributed across the four achievement levels does not reveal or disclose any information about an individual Hispanic third-grader's performance in that school. However, a statistical report for a different school shows that all of the Hispanic third-graders scored below the target performance level of proficient, and all of the White third-graders scored at or above the proficient level. The report for the second school reveals or discloses information about the performance of both White and Hispanic third-graders in this school—specifically, that each of the White third-graders reached or exceeded the performance target, while each of the Hispanic third-graders in the school failed to reach the target performance level. This release results in an attribute disclosure since specific performance can be associated with all of the students in clearly definable subgroups. (Preventing this type of inadvertent disclosure is the focus of a companion SLDS Technical Brief, *Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting.*)

Protecting Confidentiality Through De-Identification and Anonymization of Data

Other terms that are used in discussing confidentiality include de-identification and anonymization. These concepts are central to protecting against disclosures in data files that are shared with external education researchers. The term *de-identified* information is used to describe records that have enough personally identifiable information removed or obscured such that the remaining information does not identify an individual and there is no reasonable basis to believe that the information can be used to identify an individual. The FERPA 2008 regulations subsection on de-identified records allows for the nonconsensual release of student level information from education records, provided that (1) all personally identifiable information is removed and (2) there is a reasonable determination that a student's identity is not personally identifiable. In making this determination, both single and multiple data releases from the education records should be taken into account along with other information available from other sources (34 CFR § 99.31(b)(1)).

In the 2008 issuance of the Final Rule for revisions to the FERPA regulations, the Department of Education referred interested parties to the Federal Committee on Statistical Methodology's Statistical Policy Working Paper 22 *Report on Statistical Disclosure Limitation Methodology* for advice on ways to de-identify student level data (73 Fed. Reg. 74806-35, Dec 9, 2008). The Working Paper includes techniques that can be used to protect against disclosures in student level records in a data file as well as techniques that can be used to protect against disclosures in aggregate tabular reports.

Techniques described that can be used to protect student level data include generalizing the data by grouping continuous values and applying top and bottom coding to either continuous or categorical data to avoid outliers; suppressing the data by deleting entire records or parts of records; introducing "noise" into the data by adding small amounts of variation into selected data; swapping the data by exchanging certain data elements in one record with the same data elements from a similar record; blanking and imputing for randomly selected records; and blurring by replacing data with the average value by replacing a selected value (e.g., an outlier) of a data element with the average value for that data element for the entire group.

Techniques described to avoid disclosures in aggregate tabular data include establishing minimum cell sizes, suppression, complementary suppression, random rounding, controlled rounding, controlled tabular adjustment, and special rules to protect against disclosures that might include additional restrictions on publishing such as requiring results on more than one cell in a distribution, requiring certain size categories, and collapsing across categories.

Once a data file is de-identified, the FERPA regulations indicate that a re-identification code may be attached to the data file so that the file can be released for use for education research (34 CFR § 99.31(b)(2)). While the de-identified data file with a re-identification code does not provide external researchers with personally identifiable information about students, a researcher is able to return to the source that issued the data to request additional data elements that can be added using the re-identification code.

The re-identification code should be independent of any of the personally identifiable information. Only a limited number of staff should have knowledge of the method used to produce the code. Under FERPA, the re-identification code (1) may not be used for any purpose other than matching the de-identified records to the source to obtain additional information for education research; (2) may not be used to identify a student or personally identifiable information about a student; and (3) may not be based on a student's Social Security Number or other personal information (34 CFR § 99.31(b)(2)).

To understand how this would work, take the case of a school district that received a data request from an external researcher who is interested in analyzing academic gains for students who participated in an afterschool enrichment program. To do this, the district creates a fully de-identified data file that includes the relevant individual student records drawn over the researcher-specified time period for a subset of data elements that do not identify individual students. During the course of the analysis, the researcher discovers that several additional data elements and an additional year of data are needed to produce a robust analysis. The district data manager uses the re-identification codes to identify the same set of students and create an extract file that includes the additional data elements and an additional year of data for those students. The researcher uses the code to link the new data to the existing analysis file and proceeds with the analysis.

Anonymization takes the data one step beyond de-identification. That is, anonymized data are data that have been de-identified, *and* they do not include a re-identification code. In an anonymized data file, the student case numbers in the data records cannot be linked back to the original student record system. Returning to the examples discussed above, anonymized data would not be useful to staff using data to monitor the progress and performance of individual students. However, if a professor at a university reads the research report from the analysis of academic gains of students in the afterschool enrichment program and decides that he or she would like to have a class of graduate students apply different analytic procedures to see if the results can be replicated, an anonymized file could be produced from the de-identified file used by the researchers to serve this purpose. To do this, the re-identification code must be removed and the file should be reviewed to ensure that additional statistical disclosure techniques do not need to be applied. The documentation for the anonymized data file should identify any disclosure limitation techniques that were applied and their implications for the analysis.

Data Stewardship and Privacy Framework

Maintaining personally identifiable information in student education records carries both legal and ethical responsibilities for protecting the information and for ensuring the proper handling and use of the information. These concepts are part of data stewardship. The American Statistical Association's Committee on Privacy and Confidentiality cites the U.S. Census Bureau definition of data stewardship as an "organizational commitment to ensure that identifiable information is collected, maintained, used, and disseminated in a way that respects privacy, ensures confidentiality and security, reduces reporting burden, and promotes access to statistical data for public policy."

These elements of data stewardship are enacted in the various federal privacy and confidentiality laws that govern the use of personally identifiable information, including the Privacy Act of 1974, the Paper Work Reduction Act of 1980, the FERPA of 1974, the Education Sciences Reform Act of 2002 (and related authorizing laws from 1988 and 1994), and the Confidential Information Protection and Statistical Efficiency Act of 2002. They are also included in a set of tenets known as Privacy Principles or, alternatively, as Fair Information Practices. These Privacy Principles, which have been credited as forming the framework for most modern privacy laws, can be traced to the 1973 U.S. Department of Health, Education, and Welfare (HEW) report *Records, Computers and the Rights of Citizens,* Report of the Secretary's Advisory committee on Automated Personal Data Systems. The 1973 report recommended the enactment of a *Federal Code of Fair Information Practice,* consisting of a set of privacy principles.

The 1973 privacy principles set the stage for the passage of three landmark pieces of legislation—the Privacy Act, FERPA, and the Paper Work Reduction Act. Parental complaints about intrusive surveys and other data-collection activities have been cited as one reason for the enactment of FERPA (The 1977 Privacy Protection Commission, Chapter 10 Record Keeping in the Education Relationship).

These privacy principles were described in the 1973 HEW Report as safeguard requirements for data systems that include personally identifiable information. Each of these principles can be found in FERPA and the FERPA Regulations. The first principle, that there should be no secret records of personal data, is evident in the required FERPA annual notification to parents and eligible students of their right to inspect and review the student's education records (20 U.S.C. § 1232g (e); 34 CFR § 99.7). The second principle, that an individual has the right to know what personal information is retained and how it is used, is operationalized through the right to inspect and review the student's education record (20 U.S.C. § 1232g (e); 34 CFR § 99.10) and through the permissible uses of the information which are described in 20 U.S.C. § 1232g (b) and 34 CFR § 99.31. The third principle, the limitation of alternative uses of personal information without consent, is evident in the FERPA requirement that the parent or eligible student provide written consent for the student's information to be used for any purpose not specified in law (20 U.S.C. § 1232g (b)(1); 34 CFR § 99.30). The fourth principle, that an individual has the right to correct or amend a record of personal information, is addressed in law through the requirement that the parent or

The 1973 Fair Information Practices included five principles:

1. There must be no personal data record keeping systems whose very existence is secret.
2. There must be a way for an individual to find out what information about him or her is in a record and how it is used.
3. There must be a way for an individual to prevent information about him that was obtained for one use from being used or made available for other purposes without his consent.
4. There must be a way for an individual to correct or amend a record of identifiable information about him.
5. Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for the intended use and must prevent misuse of the data.

an eligible student be provided an opportunity to challenge and seek a correction of the content of a student's record "to insure that the records are not inaccurate, misleading, or otherwise in violation of the student's privacy" (20 U.S.C. § 1232g (a)(2); 34 CFR § 99.20). Finally, the fifth principle, the obligation to prevent the misuse of any personal data maintained and to ensure the reliability of the data for the intended use, is codified in the limitations on permissible uses of the information (20 U.S.C. § 1232g (b)(1)(A-J); 34 CFR §§ 99.31, 99.33–99.35).

The Privacy Act of 1974 called for a Commission to evaluate the implementation of the Privacy Act. The resulting Privacy Protection Study Commission expanded the HEW Commission's list of five Fair Information Practices to a set of eight principles, variations of which have been adopted broadly nationally and internationally.

The recent **Department of Homeland Security and Chief Information Officer Fair Information Practice Principles** include the following:

1. **TRANSPARENCY—**providing notice to the individual regarding the collection, use, dissemination, and maintenance of personally identifiable information.

2. **INDIVIDUAL PARTICIPATION AND REDRESS—**involving the individual in the process of using personally identifiable information and seeking individual consent for the collection, use, dissemination, and maintenance of personally identifiable information. Providing mechanisms for appropriate access, correction, and redress regarding the use of personally identifiable information.

3. **PURPOSE SPECIFICATION—**specifically articulating the authority that permits the collection of personally identifiable information and specifically articulating the purpose or purposes for which the personally identifiable information is intended to be used.

4. **DATA MINIMIZATION AND RETENTION—**only collecting personally identifiable information that is directly relevant and necessary to accomplish the specified purpose(s). Only retaining personally identifiable information for as long as is necessary to fulfill the specified purpose(s).

5. **USE LIMITATION—**using personally identifiable information solely for the purpose(s) specified in the public notice. Sharing information should be for a purpose compatible with the purpose for which the information was collected.

6. **DATA QUALITY AND INTEGRITY—**ensuring, to the greatest extent possible, that personally identifiable information is accurate, relevant, timely, and complete for the purposes for which it is to be used, as identified in the public notice.

7. **SECURITY—**protecting personally identifiable information (in all media) through appropriate administrative, technical, and physical security safeguards against risks such as loss, unauthorized access or use, destruction, modification, or unintended or inappropriate disclosure.

8. **ACCOUNTING AND AUDITING—**providing accountability for compliance with all applicable privacy protection requirements. Including all identified authorities and established policies and procedures that govern the collection, use, dissemination, and maintenance of personally identifiable information. Auditing for the actual use of personally identifiable information to demonstrate compliance with established privacy controls.

SLDS Technical Briefs on Privacy

The principles of the Fair Information Practice Principles provide the framework for a sound privacy and confidentiality data protection program. The information practices are recurring themes in each of the reports in the NCES series of Technical Briefs on various aspects of the protection of personally identifiable information in students' education records. In addition to this Technical Brief, the Briefs include the following topics.

Data Stewardship

An understanding of the principles of data stewardship at the school, district, and state levels provides an essential foundation for ensuring student privacy. Data stewardship starts with decisions as to what personally identifiable information is needed to successfully monitor each student's progress through the education system. Data stewardship also involves a commitment to ensuring that personally identifiable information is collected, maintained, used, and disseminated in a way that respects privacy, ensures confidentiality and security, and promotes access to data for policy formation. To provide data stewardship, there is a need for clearly established policies and procedures that govern collection, storage, processing, and access to an individual student's education records. Role-based or managed access to individual data is one key element of data stewardship. Specifically, policies and procedures should identify who within an educational agency or institution is authorized to access the records and the conditions under which they may be accessed and released. Policies could cover topics such as limiting access to "need to know" and rules and procedures to prohibit authorized users from looking at information they are not authorized to access (i.e., browsing). Procedures could include the use of signed statements of nondisclosure for authorized staff, specified methods for access and retrieval of individual records, and the identification of a secure location for their use (Fair Information Practice Principles 1 through 8).

Electronic Data Security

The development and maintenance of an efficient state longitudinal data system requires the use of an electronic record system. Because these data systems include personally identifiable student information, they should be in an electronically secure environment. All data and the hardware, software, and network infrastructure should be firewall secure and password protected to be safe from unauthorized external access. Furthermore, electronic encryption or secure networks should be used to transmit data with personally identifiable information between different entities (e.g., between the district and the state agencies) (Fair Information Principles 7 and 8).

Statistical Methods for Data Protection in Aggregate Reporting

Using information contained in student education records and related state longitudinal data systems for reporting and research requires reporting information on aggregates of students. Such reporting requires the identification and use of appropriate disclosure avoidance techniques to protect the identity of individual students in publicly available information. Because education data are reported at multiple levels (i.e., school, district, state, and federal) and in external studies, care must be taken to avoid inadvertent disclosures that can occur through comparisons of released data across reporting levels. Some current practices can result in inadvertent disclosures. A set of reporting rules that offers one approach to protecting identifiable student information in aggregate reported data will be presented. The goal of these reporting rules is to have an easy to understand and implement set of steps that can be used to protect personally identifiable student data in aggregate data. To facilitate the implementation of these rules, NCES will also provide an online tool that can be used to implement the rules (Fair Information Principle 7).

External Data Use and Written Agreements

The FERPA regulations include provisions that permit the nonconsensual release of de-identified data sets with re-identification codes to facilitate external research (34 CFR § 99.31(b)). Each of these concepts is discussed. In addition, the FERPA regulations include provisions that permit state and local educational authorities to redisclose personally identifiable information from education records to organizations conducting studies pursuant to the terms of

34 C.F.R. § 99.31(a)(6)(ii)(C) through the use of written agreements that identify and codify the terms of data sharing. Recommended components of these agreements will be discussed and a model template for an agreement will be provided (Fair Information Principles 5, 6, and 7).

Training

To successfully implement a privacy and confidentiality program for student education records, the managers of student education record systems should provide relevant staff at the state, district and school levels with periodic training to inform them of the continuing data use and data protection provisions in FERPA and other applicable privacy and security statutes and to train them on methods for compliance. These training needs are identified, with suggestions for specific content, in the guidance documents.

Data stewards and analysts will need training on newly identified disclosure limitation procedures and reporting rules for the increased protection of personally identifiable information in student education records. The technology staff should be trained on secure data transmissions, and data stewards and data managers should be trained on internal access rules and procedures and on the use of written data agreements and signed statements of nondisclosure (All Fair Information Principles, but especially 5, 6, 7, and 8)

Summary

This series of SLDS Technical Briefs is intended to open a conversation with education practitioners responsible for developing and using electronic student record systems about student privacy considerations that arise in these record systems.

NCES welcomes input on this and each of the forthcoming SLDS Technical Briefs on Privacy. You may direct comments to SLDStechbrief@ed.gov.

References

American Statistical Association, Committee on Privacy and Confidentiality. *Key Terms/ Definitions in Privacy and Confidentiality.* Alexandria, VA: Retrieved from http://www.amstat.org/committees/pc/keyterms.html on 6/17/2010.

Code of Federal Regulations, Title 34—Education, Part 99. *Family Educational and Privacy Rights,* (34CFR99). Washington, DC: GPO Access e-CFR. Retrieved from http://ecfr.gpoaccess.gov/cgi/t/text/ext-idx?c=ecfr&sid=44d350c26fb9cba4a156bf805f297c9e&tpl=/ecfrbrowse/Title34/34cfr99_main_02.tpl on 9/9/2010.

Duncan, George T., Jabine, Thomas B. and de Wolf, Virginia A., Editors (1993). *Private Lives and Public Policies: Confidentiality and Accessibility of Government Statistics.* Panel on Confidentiality and Data Access, National Research Council. Washington, DC: National Academy Press.

Federal Register, Office of Management and Budget. *Implementation Guidance for Title V of the E-Government Act, Confidential Information Protection and Statistical Efficiency Act of 2002* (CIPSEA). Washington, DC: Vol. 72, No. 115/ Friday, June 15, 2007. Retrieved from http://www.whitehouse.gov/sites/default/files/omb/assets/omb/fedreg/2007/061507_cipsea_guidance.pdf on 9/9/2010.

Federal Register, Part II Department of Education. *Family Educational Rights and Privacy; Proposed Rule (34 CFR Part 99).* Washington, DC: Vol. 73, No. 57/ Monday, March 24, 2008.

Federal Register, Part II Department of Education. *Family Educational Rights and Privacy; Final Rule (34 CFR Part 99).* Washington, DC: Vol. 73, No. 237/ Tuesday, December 9, 2008.

McCallister, E., Grance, T., and Scarfone, K. (2010). *Guide to Protecting the Confidentiality of Personally Identifiable Information (PII): Recommendations of the National Institute of Standards and Technology* (NIST Special Publication 800-122). National Institute of Standards and Technology, U.S. Department of Commerce. Washington, DC: Retrieved from http://csrc.nist.gov/publications/nistpubs/800-122/sp800-122.pdf on 5/4/2010.

National Academy of Sciences, Committee on National Statistics and the Center for Education, Workshop (2009). *Protecting Student Privacy and Facilitating Education Research.* Washington, DC: National Academy Press.

National Academy of Sciences, Institute of Medicine, Committee on Health Research and the Privacy of Health Information: The HIPAA Privacy Rule (2009). *Beyond the HIPAA Privacy Rule: Enhancing Privacy, Improving Health Through Research.* Washington, DC: National Academy Press.

Office of Management and Budget, Federal Committee on Statistical Methodology, (2005). Statistical Policy Working Paper 22, *Report on Statistical Disclosure Limitation Methodology.* Retrieved from http://www.fcsm.gov/working-papers/spwp22.html on 9/9/2010.

Office of Management and Budget. OMB Memorandum M-03-22, *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002.* Retrieved from http://www.whitehouse.gov/omb/memoranda_m03-22/ on 9/9/2010.

Office of Management and Budget. OMB Memorandum M-07-16, *Safeguarding Against and Responding to the Breach of Personally Identifiable Information.* Retrieved from http://www.whitehouse.gov/sites/default/files/omb/assets/omb/memoranda/fy2007/m07-16.pdf on 9/9/2010.

Public Law 107-347, E-Government Act of 2002, Title II, Sec. 208 (d). *Privacy Provisions.* Washington, DC: GPO Access. Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=107_cong_public_laws&docid=f:publ347.107.

Public Law 111-05, American Recovery and Reinvestment Act of 2009, Title VIII—Departments of Labor, Health and Human Services, and Education, and Related Agencies, Institute of Education sciences, Stat. 184, Washington DC: GPO Access. Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_public_laws&docid=f:publ005.111 on 9/9/2010.

U.S. Code, Title 5—Government Organization and Employees, Chapter 5—Administrative Procedure, Subchapter II—Administrative Procedure, Section 552a. *Records Maintained on Individuals (Privacy Act),* (5USC522a). Washington, DC: GPO Access. Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/usc.cgi?ACTION=RETRIEVE&FILE=$$xa$$busc5.wais&start=312761&SIZE=77292&TYPE=TEXT on 9/9/2010.

U.S. Code, Title 20—Education, Chapter 31—General Provisions Concerning Education, Subchapter III—General Requirements and Conditions Concerning Operation and Administration of Education Programs: General Authority of Secretary, Part 4—Records, Privacy, Limitation on Withholding Federal funds, Section 1232g. *Family Educational and Privacy Rights,* (20USC1232g). Washington, DC: GPO Access. Retrieved from http://frwebgate4.access.gpo.gov/cgi-bin/TEXTgate.cgi?WAISdocID=799486197532+0+1+0&WAISaction=retrieve on 9/9/2010.

U.S. Code, Title 44—Public Printing and Documents, Chapter 35—Coordination of Federal Information Policy, Subchapter I—Federal Information Policy, Section 3512. *Public Protection (Paperwork Reduction Act).* (44USC3512). Washington, DC: GPO Access. Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/usc.cgi?ACTION=RETRIEVE&FILE=$$xa$$busc44.wais&start=978642&SIZE=2355&TYPE=TEXT on 9/9/2010.

U.S. Code, Title 20—Education, Chapter 70—Strengthening and Improvement of Elementary and Secondary Schools, Subchapter I—Improving the Academic Achievement of the Disadvantaged, Part A—Improving Basic Programs Operated by Local Educational Agencies, Subpart 1—Basic Program Requirements, Section 6311. *State Plans,* (20USC6311). Washington, DC: GPO Access. Retrieved from http://frwebgate2.access.gpo.gov/cgi-bin/TEXTgate.cgi?WAISdocID=bULwJH/21/1/0&WAISaction=retrieve on 9/9/2010.

U.S. Code, Title 20—Education, Chapter 76—Education Research, Statistics, Evaluation, Information, and Dissemination, Subchapter I—Education Sciences Reform, Section 9573. *Confidentiality,* (20USC9573). Washington, DC: GPO Access. Retrieved from http://frwebgate.access.gpo.gov/cgi-bin/usc.cgi?ACTION=RETRIEVE&FILE=$$xa$$busc20.wais&start=10326022&SIZE=10154&TYPE=TEXT on 9/9/2010.

U.S. Code, Title 20—Education, Chapter 76—Education Research, Statistics, Evaluation, Information, and Dissemination, Subchapter II—Educational Technical Assistance, Section 9607. *Grant Program for Statewide, Longitudinal Data Systems,* (20USC9607). Washington, DC: GPO Access. Retrieved from http://frwebgate3.access.gpo.gov/cgi-bin/TEXTgate.cgi?WAISdocID=FKr6BA/0/1/0&WAISaction=retrieve on 9/9/2010.

U.S. Department of Health and Human Services, Report of the HEW Secretary's Advisory Committee on Automated Personal Data Systems (1973). *Records, Computers and the Rights of Citizens.* Washington, DC: Retrieved from http://aspe.hhs.gov/datacncl/1973privacy/tocprefacemembers.htm on 5/11/2010.

U.S. Privacy Protection Study Commission (1977). *Personal Privacy in an Information Society.* Retrieved from http://epic.org/privacy/ppsc1977report/c1.htm on 9/9/2010.