Comments of the World Privacy Forum regarding Notice of Proposed Rulemaking, Family Educational Rights and Privacy, 34 CFR Part 99, RIN 1855-AA05, Docket ID ED-2008-OPEPD-0002

To:

LeRoy S. Rooker
U.S. Department of Education
400 Maryland Avenue S.W.
Room 6W243
Washington, DC 20202-5020

Via Federal eRulemaking Portal

May 6, 2008

The World Privacy Forum appreciates the opportunity to comment on the Department's proposed rule change to the Family Educational Rights and Privacy Act at 73 Federal Register 15574 (March 24, 2008). The World Privacy Forum is a non-partisan, non-profit public interest research and consumer education organization. Our focus is on conducting in-depth research and analysis of privacy issues, in particular issues related to information privacy. More information about the activities of the World Privacy Forum is available at <http://www.worldprivacyforum.org>.

Our comments focus on several aspects of the Notice of Proposed Rulemaking (NPRM), notably, the definition and handling of directory information and personally identifiable information. We also comment on the use of full tax returns to determine eligibility. And finally, we comment on the issue of outsourcing, including the need for audit trails in regards to the proposed expansion of the *school official* exemption.

I. Definitions

A. Directory Information

WPF supports exclusion of Social Security Numbers in definition

The proposed rule seeks to change the definition of *directory information*. In general, the World Privacy Forum supports the change and supports in particular the provision that Social Security Numbers are not included in the definition of directory information.

Disclosure of electronic identifiers is problematic

We find some problems with the proposed definitions, one of which is the proposed disclosure and handling of *electronic identifiers*.

This is the text of the new language proposed for 34 CFR § 99.3:

> *Directory information* means information contained in an education record of a student that would not generally be considered harmful or an invasion of privacy if disclosed.
>
> (a) 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.
>
> (b) Directory information does not include a student's social security number or student identification (ID) number.
>
> (c) Directory information includes a student's user ID or other unique personal identifier used by the student for purposes of accessing or communicating in electronic systems, **but only if the electronic identifier cannot be used to gain access to education records except when used in conjunction with one or more factors that authenticate the user's identity, such as a personal identification number (PIN), password, or other factor known or possessed only by the authorized user.**

We highlight the last portion because we find it troublesome that a student's electronic identifier can be made public. The explanation on page 15575 for this change is the following:

> Much of the directory-based software used for these systems, as well as protocols for electronic collaboration by students and teachers within and among institutions, essentially cannot function without making an individual's user ID or other electronic identifier publicly available in these kinds of systems.

We recognize that there may be a better reason than this for making electronic identifiers public, but the stated reason appears inadequate.

It is not apparent why disclosure <u>to the world</u> of a student's electronic identifier is an essential feature of electronic collaboration systems. We can envision how some sharing of an electronic identifier within a class or for other uses internal to the school or networked schools might be appropriate. But it is unreasonable to state that absolutely everyone, even those outside the school system who are not using any electronic collaboration, need the electronic identifier. Similar technologies in other sectors such as finance and health have been able to largely

eliminate this problem. We see no reason why this problem cannot be eliminated from this sector in a similar fashion, particularly when the electronic identifier may offer a direct route of access to other sensitive information, acting as a "key" of sorts to the entire student file.

Passwords alone do not constitute robust or even adequate security for records

An additional and significant problem is the reliance on a PIN or password as protection for the records. We reject the notion that a password or PIN will protect students and their records. We especially reject the notion that a student's electronic identifier will not be abused if "protected" with a PIN or a password.

First, every significant study of passwords shows that many people use either an easily guessed password or they use and reuse the same password for everything. These kinds of common insecure password usage patterns, in fact, form the basis of many phishing and other security attacks that seek to steal individuals' identity information and have been successful at doing so in many cases. The security literature is quite clear and consistent in this area, which is one of the reasons two-factor authentication is overtaking single factor (password/PIN) authentication as the baseline norm.[1]

The Federal Trade Commission has set out significant guidance to business and consumers in the area of passwords.[2] Additionally, the GAO has regularly given testimony regarding insecure Federal Agency systems, and has noted an over-reliance on passwords as a significant security problem. In its most recent testimony regarding this issue from March 2008, the GAO stated:

> …another IG reported that the agency security awareness program needs to increase employees' awareness of social engineering techniques and the importance of protecting their usernames and passwords as a result of successful social engineering attempts.[3]

The proposed reliance on passwords in the NPRM to protect sensitive student information appears to violate the commonsense guidance routinely given in this area by the GAO, the FTC, and Federal agencies such as the FDIC and others.

We have no doubt that today's students are particularly adept at identifying the passwords of their peers. Identity thieves and others will have no difficulty cracking the average password if

---

[1] *See* Matthias Schonlau, William DuMouchel, Wen-Hua Ju, Alan F. Karr, Martin Theus and Yehuda Vardi, *Computer Intrusion: Detecting Masquerades*. Statistical Science, Vol. 16, No. 1 (Feb., 2001), pp. 58-74. *See also*: Robert Sikorski and Richard Peters, *Pacified Passwords*, Science, New Series, Vol. 278, No. 5346 (Dec. 19, 1997), pp. 2145-2146. *See also*: J. V. Hansen, *Internet Commerce Security: Issues and Models for Control Checking*, The Journal of the Operational Research Society, Vol. 52, No. 10 (Oct., 2001), pp. 1159-1164. *See also*: Detmar W. Straub and Richard J. Welke, *Coping with Systems Risk: Security Planning Models for Management Decision Making*, MIS Quarterly, Vol. 22, No. 4 (Dec., 1998), pp. 441-469.

[2] Federal Trade Commission, *Protecting Personal Information, A Guide for Business*, <http://www.ftc.gov/bcp/edu/pubs/business/privacy/bus69.pdf>.

[3] GAO, *Information Security: Progress Reported, but Weaknesses at Federal Agencies Persist*, March 12 2008, GAO-08-571T. < http://www.gao.gov/new.items/d08571t.pdf>.

they think that they can obtain valuable records. The Department should not assume that an ordinary password would protect academic records, student financial records, and other education records. We suggest that there has to be a better solution to whatever the real problem is here.

Student electronic identifiers should not be part of the directory information. Password or PIN protection is not enough to prevent abuse, particularly in connection with the student electronic identifier – the key to many records.

The need to establish categories of directory information

Part of the difficulty here is the treatment of all types of directory information as the same. It is one thing to circulate a student list to parents in the school. It is something else to circulate a full list of every permissible element of directory information to the world outside the school. We suggest that the Department consider establishing categories of directory information.

Some information would be eligible for circulation within the school community, while other information might be eligible for broader circulation. We worry that administrative convenience or regulatory uncertainty may result in schools putting more information into a public directory than is really needed.

While we recognize that the statute allows student directories and that directories can serve useful purposes, a directory is still a major imposition on the privacy of a student and parent. This conclusion is even more important in this era of international identity theft activities than it was when FERPA first became law. The contents of a student directory should be based to some extent on the need-to-know principle.

Schools should be prohibited from penalizing students who opt-out

We are even more troubled by this further explanation at page 15576:

> Note that eligible students and parents have a right under FERPA to opt out of directory information disclosures and refuse to allow the student's e-mail address, user ID or other electronic identifier disclosed as directory information (except as provided in proposed § 99.37(c), discussed elsewhere in this document). This is similar to a decision not to participate in an institution's paper-based student directory, yearbook, commencement program, etc. In these cases, the student or parent will not be able to take advantage of the services, **such as portals for class registration, academic records, etc., provided solely through the electronic communications or software that require public disclosure of the student's unique electronic identifier.**

The highlighted part suggests that a school may be able to deny benefits, services, or even <u>required activities</u> to students who have exercised the right to opt-out of directory information.

Each student and parent should be able to exercise the right to opt-out without official cost or consequence. Yet the Department's only explanation here suggests that it would be okay for a school to effectively blackmail a student into forgoing an opt-out lest the student lose an important right, benefit, or opportunity.

The Department should not allow a school to impose any negative consequences on those who exercise an opt-out. A weaker alternative is for schools to be required to explain in detail the consequences of the exercise of an opt-out for each element of information included in a directory. The Department can do better here.

B. Personally Identifiable Information

We appreciate and applaud the Department's recognition of the increasing ease of identifying individuals from the use of non-identifiable information. The discussion and citations in the draft rule are generally useful and appropriate. We are troubled, however, by this language that the Department proposes to add to the definition of *Personally Identifiable Information*:

> (f) Other information that, alone or in combination, is linked or linkable to a specific student that would allow **a reasonable person in the school or its community**, who does not have personal knowledge of the relevant circumstances, to identify the student with reasonable certainty; [emphasis added]

Detailed personal information on most Americans is maintained by so many different public, private, and commercial institutions that it would be impossible to compile a list of all the unregulated sources of personal information. The buying and selling of personal information by lawful commercial data brokers goes on at a pace that would shock most consumers. If consumers knew the level of detail of that information, they would be even more troubled.

The Department wants to assess the possibility of re-identification by focusing on the abilities of a ***reasonable person in the school or its community***. That focus is far too narrow, and as such is an inappropriate guide. Many legitimate local, national, and international commercial entities – let alone identity thieves and other crooks – have the capability of taking seemingly de-identified data and re-identifying it. It is much easier to do than most people realize. For example, most Americans (87%) can be uniquely identified from a birth date, five-digit zip code, and gender.[4]

Determining whether data can be re-identified based on the capabilities of people in a single school or local community is naïve and inappropriate. Information can be transferred around the world in an instant, and anyone with even mild technical skills may be able to re-identify the data using the wealth of personal information and software tools readily available commercially, and increasingly, for free. Schools may be local institutions, but de-identified data about students can be re-identified anywhere around the globe and can be exploited by many.

---

[4] B. Malin, L. Sweeney, and E. Newton. *Trail Re-identification: Learning Who You are From Where You Have Been.* Carnegie Mellon University, School of Computer Science, Data Privacy Laboratory Technical Report, LIDAP-WP12. Pittsburgh: February 2003. *See also*: Malin, B. *Betrayed by My Shadow: Learning Data Identity via Trail Matching*. Journal of Privacy Technology. 2005; 20050609001.

Re-identification is a problem for any de-identified data, but it is much worse for student data. Why? The reason is that FERPA allows for the regular publication of student directories containing a wealth of personal information, including address and birth date. The availability of these directories – year after year for each school – makes it much easier to use existing tools and data to re-identify data. Just to make a point, you will not find a hospital that publishes a directory of in-patients or a bank that publishes a directory of credit card holders. Schools that publish student directories create a resource that will greatly assist anyone seeking to re-identify data about students and their families in both a school context and in many other contexts that extend far beyond the school setting.

At a minimum, the Department should amend the language in paragraph (f) to recognize the possibility that <u>any person</u> with appropriate knowledge and expertise may be able to re-identify the data. We suggest amending paragraph (f) thusly:

> (f) Other information that, alone or in combination, is linked or linkable to a specific individual or family that would allow (1) a reasonable person in the school or its community, who does not have personal knowledge of the relevant circumstances, to identify any individual or family with reasonable certainty; or (2) any skilled person with appropriate knowledge, expertise, and resources using currently available commercial, public, or privately-held data to reasonably infer the identity of any individual or family by either direct or indirect means.

Note that some of the language we suggest above for subparagraph (2) was borrowed from section 502(4) of the Confidential Information Protection and Statistical Efficiency Act of 2002 (CIPSEA). Note also that in addition to adding a new subparagraph, we also suggest striking the word *student* in the first subparagraph and inserting *individual or family*. It may be possible to use de-identified data to identify a parent or household and not just a student.

Establishing a firm standard for de-identified data is a difficult task, but it is nearly impossible to be too strict. The amount of identifiable personal data available increases daily, and the computer power that can be employed to re-identify data increases even faster. We have all watched the rise of identity theft as a worldwide criminal enterprise over the past decade. The problem is not going away. The Department needs to be certain that a skilled student with a criminal or mischievous bent cannot create a program that marries seemingly de-identified educational research, school directories, and commercial databases to create a new resource for identity thieves, marketers, and others.

II. Eligible Students and the Use of Tax Returns for Determining Status

The discussion of an *eligible student* brought our attention to the existing provision in § 99.31(a)(8) that uses the tax code's definition of *dependent*. On page 15578, we find this comment:

> (Institutions must first determine that a parent has claimed the student as a dependent on the parent's Federal income tax return. Institutions can determine that a parent claimed a student as a dependent by asking the parent to submit a copy of the parent's most recent Federal tax return. Institutions can also rely on a student's assertion that he or she is not a dependent unless the parent provides contrary evidence.)

We observe that a tax return is among perhaps the most sensitive financial documents for nearly every individual and family. It is offensive for the Department to encourage schools to ask parents to turn over a tax return to prove facts that can, in most cases, be readily ascertained by talking to a parent and student. Would the Department expect Warren Buffet or Bill Gates to turn over a tax return for this purpose? What about asking the faculty of a school to turn over their tax returns to the school or school system that employs them?

Any school that maintains a file of tax returns of other people for this trivial purpose will eventually trigger a scandal of front-page proportions, whether from snooping, data breach, identity theft, or just plain negligence. Most states' data breach statutes would be triggered by a breach of tax return information at a school, causing the school to need to notify all affected parties and potentially provide a year of free credit monitoring. This would be a heavy burden for many schools. This risk could be greatly mitigated by reducing this information collection in the first place.

The Department should:

(1) prohibit schools from seeking or accepting a tax return for this purpose unless other means of determining dependency have been exhausted;

(2) instruct schools to seek only the front page of a Form 1040 that shows names of dependents;

(3) allow parents to redact all numbers that appear on that page; and/or

(4) require schools to destroy any tax form submitted within 24 hours after receipt.

We specifically encourage the Department to pay attention to the privacy of parents and not just students.

III. Outsourcing and the need for audit trails

The proposed regulation in § 99.31(a)(1)(i)(B) seeks to expand the *school official* exception to include contractors, consultants, volunteers, and other outside parties to whom an educational agency or institution has outsourced institutional services or functions that it would otherwise use employees to perform.

If we correctly understand the effect of this proposal, disclosures to contractors and others would not be covered by the existing requirement in § 99.32 for maintaining a record of each disclosure. We object to allowing outsourcing disclosures without a record that will serve as an audit trail. An annual FERPA notification is not a substitute that will allow a student or parent to learn that a school disclosed a particular record and for what purpose. An audit trail is essential to protect the student and to protect the school as well. Otherwise, it may be impossible to trace who is responsible for non-compliance with FERPA, for a data breach, or for other improper activities.

We observe that the federal health privacy rule issued under the Health Insurance Portability and Accountability Act (HIPAA) requires the maintenance of audit trails (called an *accounting of disclosures*) for many disclosures to business associates and many other recipients. If HIPAA-covered entities can maintain audit trails, so can schools. We also observe that the Department of Education itself is required to maintain an accounting for all external disclosures under the Privacy Act of 1974.

The requirement for an audit trail should be universal for schools outsourcing functions that involve the transfer of student data. Audit trails are simple and inexpensive to maintain for any computerized disclosure of records. It is good administrative practice (and good liability protection too) for modern computer systems to track all disclosures (and all accesses as well). However, if the Department is not willing to require the maintenance of more audit trails, we have alternatives to suggest.

Audit trails are a common feature of modern computer systems. If maintaining an audit trail is viewed as too cumbersome or expensive for paper records, then the requirement can be limited to disclosures from computer systems. If existing computer systems do not have audit trail capability, then the Department should require that any new computer system acquired more than one year after the effective date of the new regulation must have the capability of maintaining an audit trail for outsourcing. We also encourage the Department to tell schools that they will eventually need to maintain audit trails for all uses as well as disclosures.

Finally, for any school that has an audit trail for any disclosure whether or not the audit trail is required by the rule, the Department should require the school to make the audit trail available to the student or parent. If a school already has an audit trail, there is no justification for using the rule as an excuse for denying access to a record that exists and can be easily retrieved. We encourage the Department to direct schools to make all existing audit trails available to students and parents without regard to any requirement in the rules to maintain them. Only those audit trails specifically exempt from access should be protected.

Respectfully submitted,

Pam Dixon
Executive Director,
World Privacy Forum