United States House of Representatives
Committee on Education and Labor

Hearing on
"How Data Can be Used to Inform Educational Outcomes"
April 14, 2010

Statement of Joel R. Reidenberg
Professor of Law and Founding Academic Director
Center on Law and Information Policy
Fordham University School of Law
New York, NY

Good morning Mr. Chairman, Ranking Member, and distinguished members of the Committee. I would like to thank you for the invitation to testify today and to commend you for recognizing the importance of privacy protections in the development of databases of children's educational records.

My name is Joel Reidenberg. I am a Professor of Law and the Academic Director of the Center on Law and Information Policy ("CLIP") at the Fordham University School of Law. As an academic, I have written and lectured extensively on data privacy law and policy. Of relevance to today's hearing, I directed with Jamela Debelak, CLIP's Executive Director, the CLIP report "Children's Educational Records and Privacy: A Study of Elementary and Secondary School State Reporting Systems" (Oct. 28, 2009) <http://law.fordham.edu/childrensprivacy>. I am a former chair of the Association of American Law School's Section on Defamation and Privacy and have served as an expert adviser on data privacy issues for the Federal Trade Commission, the European Commission and during the 103rd and 104th Congresses for the Office of Technology Assessment. I have also served as a Special Assistant Attorney General for the State of Washington in connection with privacy litigation. In appearing today, I am testifying as an academic expert and my views should not be attributed to any organization with which I am affiliated.

My testimony today draws on the Fordham study and I would like to make three points directly from it:

1. **States are warehousing sensitive information about identifiable children.**

2. **The Fordham CLIP study documents that privacy protections are lacking and rules need to be developed and implemented to assure that children's educational records are adequately protected.**

3. As part of basic privacy standards, strong data security is necessary to minimize the risks of data invasions, scandals and melt-downs from centralized databases of children's personal information.

My research focus on the treatment of K-12 educational records began in October 2006. As an elected member of the Millburn Township Board of Education in New Jeresey, I heard a speech by the state commissioner of education extolling the roll-out of the NJ SMART data warehouse later that fall. The NJ SMART program required our district to provide detailed, sensitive information about our school children on an identifiable basis to the state's central database. None of the commissioner's plans indicated any effort to focus data collection on truly necessary information, nor did they reflect any limitation on the purposes for use of the data once collected, nor did the plans appear to have any means for parents to check the accuracy of state-held information, and nor did the plans have any limitations on the length of storage. The only recognition that privacy might be affected by NJ SMART was an architecture that included data security mechanisms. As a Board member, I was disturbed that the state had given our district a mandate that would invade our children's privacy for ill-defined purposes in a way that appeared to put the district in clear violation of the Family Educational Rights and Privacy Act ("FERPA"). I was equally troubled that this database was established without public transparency and debate on the policy ramifications for children's privacy. Our Board and others we asked had not even heard about the program.

In delving further into the New Jersey program, it became apparent that New Jersey was part of a national trend to create state data warehouses of children's educational records driven by No Child Left Behind and more recently expanded by the American Recovery and Reinvestment Tax Act of 2009. The national trend similarly had emerged without public debate regarding privacy. As a result, we launched the Fordham CLIP study to determine what existed across the country at the state level, to assess whether states were protecting the privacy of the children's information in these databases and to make best practices and legislative reform recommendations as appropriate.

At the outset, I would like to stress that our study and I do not challenge the importance and legitimacy of data collection and use to better inform educational outcomes. Rather, I seek to highlight the critical need for policy makers to incorporate privacy rules in the planning and implementation of these systems so that the important and legitimate goals of educational accountability do not undermine privacy and so that the important and legitimate privacy concerns do not pose unnecessary obstacles to educational accountability.

1. States are warehousing children's sensitive personal information

The Fordham study found that most states have established state-wide databases of children's educational records. The information held at the state level is typically identified or identifiable to individual children because the databases use unique identifiers for each child and very few states use systems that establish a firewall to keep

the identity of individual students known only at the local level. One-third of the states track students through their social security numbers. In other words, most states are developing systems that centralize at the state level each individual child's information rather than transferring data aggregated by cohorts to the state level.

For a disturbing number of states such as Alabama, Arizona, Maryland, Nevada and Oklahoma, key information on the data warehouse programs including the types of data that were being collected and used were not publicly available. This means that state governments are conducting major data processing operations involving children's sensitive information essentially in secret from parents.

In states where information was publicly available on the data warehouse programs, the Fordham study found that states were collecting children's personal information to comply with NCLB reporting obligations such as test scores, race, ethnicity, gender, and disability status. However, the states were also collecting sensitive information well beyond NCLB reporting requirements. The following table gives some examples of the sensitive data collected by states.

Longitudinal Databases and Sensitive Data

- ***32% of states collect children's social security numbers***
- ***22% of states record student pregnancies***
- ***46% of states have a mechanism in place to track children's mental health, illness and jail sentences***
- ***72% of states collect children's family wealth indicators***

Source: Fordham CLIP Study, "Children's Educational Records and Privacy: A Study of Elementary and Secondary School State Reporting Systems" (Oct. 28, 2009), p. 27

Many additional data elements included in the state databases do not appear to be collected for NCLB reporting purpose nor for core educational assessment purposes. Louisiana schools, for example, must report to the state the social security number of each child who is disciplined for the use of foul language in school.

Data warehouses appear to gather data for other goals like the delivery of social services. For example, Florida uses social security numbers to collect information about its K-12 children and collects the birth weight of a teenage mother's baby. While the

birth weight of a teenage mother's baby can be valuable information to anticipate social service needs, the decision to include this information as part of an educational record at the state level permanently linked to the teenager and the baby raises many privacy risks that need to be justified and balanced against the actual benefits for the mother and child. The following table illustrates some of these types of data found in the state data warehouses.

Examples of Other Sensitive Data
Collected by the States

- ***Birth order***
- ***Birth weight of a student's baby***
- ***Victim of peer violence***
- ***Medical test results***
- ***Parental education level***
- ***Mental health problems***
- ***Criminal history***

Source: Fordham CLIP Study, "Children's Educational Records and Privacy: A Study of Elementary and Secondary School State Reporting Systems" (Oct. 28, 2009), p. 31

In developing data warehouses, the U.S. Department of Education has encouraged the use of interoperable data standards. Organizations, such as the Data Quality Campaign and the Standards Interoperability Framework Association, have significantly advanced the development of common data protocols. These common protocols are valuable to improve the efficiency of data collection and use. But, the use of interoperable data standards across state lines also means that the creation of a national database of children becomes a turn-key operation. Until the recent efforts of the Data Quality Campaign, basic privacy protections were not included as key components of the work on common data standards.

2. The Lack of Privacy Protection

The Fordham study showed that the state data warehouses of children's information typically lacked basic privacy protections and, often, were not in compliance with FERPA.

Existence of Key Privacy Protections

- **Only 18 states have detailed access and use restrictions**
- **Only 18 states require database users to enter into confidentiality agreements**
- **Only 10 states have data retention policies**
- **49 states make FERPA information accessible on the Internet, but for many the information is hard to find, vague or incomprehensible**

Source: Fordham CLIP Study, "Children's Educational Records and Privacy: A Study of Elementary and Secondary School State Reporting Systems" (Oct. 28, 2009), p. 39

As a starting point, the states' lack of transparency for these databases is deeply troubling. Our research team had significant difficulty and was unable to find publicly available information on the data collected by many states. As far as parents are concerned, this means that state governments have created secret surveillance systems for their children. The non-transparent nature of these systems also means that state government can avoid public accountability for its treatment of children's personal information.

The technical architectures generally did not adequately seek to de-identify children's information at the state level. To the extent that outcome assessment can effectively be accomplished by examining cohorts at the state level, rather than individual children, there is no need for the state educational agency to have individual student records. The use of truly anonymous information would avoid privacy issues. However, we did not systematically see careful attention to architectures that established identity firewalls. Professors Krish Muralidhar and Rathindra Sarathy have demonstrated that re-identification of specific children from purportedly anonymous student information is already a problem in the context of public reporting on school performance.[1]

Data minimization, a basic privacy principle that collections of personal information should not be conducted as general fishing expeditions, is absent as a guiding policy for the state warehouses. The scope of sensitive children's information that is

[1] Krish Muralidhar & Rathindra Sarathy, "Privacy Violations in Accountability Data Released to the Public by State Educational Agencies," paper presented to the Federal Committee on Statistical Methodology Research Conference, Washington DC, November 2-4, 2009 available at: <http://gatton.uky.edu/faculty/muralidhar/EdPrivacyViolation.pdf> (last visited Apr. 9, 2010).

collected by states appears to be excessive with respect to the context and core educational purposes of the databases.

The state data warehouses generally did not have clear legal limitations on the purpose for which data could be accessed and used. Without purpose limitations, states, such as New Jersey, are in facial violation of FERPA. FERPA only permits local schools to report data to state agencies in identifiable format for "audit and evaluation" purposes. The lack of purpose limitations strongly suggests that states will begin a mission creep and use children's educational data for a multiplicity of purposes unrelated to assuring the educational performance of the state's schools. Most states also did not explicitly require state officials to agree to confidentiality before accessing student information.

The states by and large ignore data retention policies. The lack of storage limits means that a child's third grade peccadillo and youthful indiscretions will indeed become a "permanent record" since states store detailed disciplinary and social information, including in some instances if a child was the victim of bullying. The lack of storage limitations is a facial violation of FERPA as FERPA requires that data transferred to state authorities for audit and evaluation purposes not be retained longer than necessary to accomplish those permissible purposes. The lack of durational limits also undermines other important public policies. For example, the detailed disciplinary information collected on identified students, including involvement and convictions under the juvenile justice system will be held indefinitely as part of the "educational records" database. While the juvenile records are typically sealed and may be expunged when a minor reaches adulthood, the state's educational database without a data retention policy does not provide any such protection.

Many states outsource the data processing services for their data warehouses. While security and confidentiality provisions can be found in some of these contracts, the clauses are typically very circumspect with respect to the vendor's obligations. Vendor contracts are generally silent with respect to uses and retention of data by the vendor.

The Fordham CLIP study identified key privacy protections that need to be implemented for children's educational record databases:

- *States should implement a technical architecture to prevent access to identifiable information beyond the school officials who need to know*
- *States that outsource data processing should have comprehensive agreements that explicitly address privacy*
- *States should limit data collection to necessary information for articulated, defined purposes*
- *States should have specific data retention policies and procedures*
- *States should explicitly provide for limited access and use of the children's data*
- *States should provide public notice of state data processing of children's information*

3. Strong data security is necessary to minimize the risks of data invasions, scandals and melt-downs from centralized databases of children's personal information.

In addition to basic privacy protections, data security is critical when information relating to identifiable children is centralized at the state level. Data security measures do not address the essential policy decisions for privacy protections like data minimization, purpose limitations, and defined storage periods. But, data security measures play a critical role in the implementation of privacy protections specifically with respect to the prevention of unauthorized access, use and disclosure of personal information.

The centralization of children's information at the state level increases the risks and scope of loss from security incidents. The centralization means that data security breaches will be on a larger scale than if data were held solely at the local level. For example, according to the Congressional Research Service up to 1.4 million residents of Colorado had their names, social security numbers and birth dates compromised when a database from the state department of human services was stolen from a private contractor in Texas.[2]

It is inevitable that security of the children's information will be compromised. The experiences in the financial services sector that have been revealed by data security breach notification laws reflect the magnitude of this risk. Despite the deployment of significant resources and the economic incentive for banks to avoid liability, the number of compromised credit cards in the United States is staggering. The Heartland Payment Systems breach alone in 2009 involved more than 100 million credit and debit card transactions. State departments of education have neither the resources nor the same high level of incentive to protect children's information to the degree that the financial services sector does.

The substantial security risks to children's educational records in data warehouses can be illustrated by a few examples:

- **Data spills** occur when school or state officials fail to assure adequate access controls and encryption for student records

[2] CRS Report for Congress, Data Security Breaches: Context and Incident Summary, p. 62 (May 7, 2007) available at: <http://www.fas.org/sgp/crs/misc/RL33199.pdf>

Recent Data Spills

Catawba County, NC: names, test scores and SSNs of school children exposed on the web (2006)
Nashville, TN: personal information of 18,000 students and 6,000 parents released on the internet from state data warehouse program (2009)
100 Public Schools and Local Government Entities: FTC warns that their files of personal information can be found freely on the web with P2P technology (2010)

- **Hackers** gain access to data when it is insufficiently protected

Hacking Cases

Churchill High School, Potomac, MD: students hacked school records system to alter data
Haddonfield High School, Haddonfield, NJ: students hacked into school records database

- **Data loss and theft** compromise educational records when they are insufficiently protected

Loss and Theft Cases

Broward County, FL: ChildNet lost personal information on adoptive and foster families including SSNs, passport numbers, credit data, drivers' license information
Chicago Public Schools, IL: lost personal information on 40,000 teachers and employees when 2 laptops stolen
Colorado: lost health records on 1,600 named, autistic children when laptop stolen from state employee's home (2005)
Greenville County School District, NC: lost personal information on 100,000 students and staff when district laptops auctioned off

- **Data spys and voyeurs** who are internal employees with access privileges abuse their access to personal information for personal gain

Spying and Voyeur Cases

UCLA Medical Center: hospital worker sells celebrity patient information to media
IRS: tax agent in Kentucky convicted for spying on 200 actors and sports figures

Strong data security for children's educational records is, thus, essential. Four critical features for a strong security system are:

- *States should avoid the storage of identifiable information whenever possible.*
- *States should use state-of-the art encryption to protect children's data*
- *States should have robust access control and use authorization policies in place*
- *States should, like the IRS, maintain audit logs that track system use to detect intrusions and police internal misuse*

Conclusion

The Fordham CLIP Study recommends several measures that I believe Congress should consider as a condition of continued federal funding of state data warehouses of children's information:

1) **Require that states articulate through statute or regulation the justification for the collection of each element of identifiable information.** This assures that the legitimate uses are transparent and sufficiently compelling to warrant the privacy trade-offs.
2) **Require that states define specific data retention limitations that are clearly linked to the specific purposes for which the data is originally collected.** This reduces the risks of data spills, protects against mission creep, and
3) **Require that states adopt an oversight mechanism for the collection and use of children's educational data.** A Chief Privacy Officer in the state departments of education would, like the CPOs in the federal Department of Homeland Security and Department of Justice, provide transparency to the public and oversight for compliance with privacy requirements.

Biography

Joel R. Reidenberg is Professor of Law and the Founding Academic Director of the Center on Law and Information Policy at Fordham Law School. He is a former Associate Vice President for Academic Affairs and Associate Chief Academic Officer of Fordham University and a former President of the University's Faculty Senate (the governing body of the university-wide faculty).

Professor Reidenberg is an expert on information technology law and policy. His published books and articles explore both information privacy law as well as the regulation of the internet. He teaches courses in Information Privacy Law, Information Technology Law, and Intellectual Property Law. He has held appointments as a visiting professor at the Université de Paris 1 (Panthéon-Sorbonne), at the Université de Paris V (René Descartes) and at AT&T Laboratories - Public Policy Research .

Professor Reidenberg has served as an expert adviser on data privacy matters for the U.S. Congress, the Federal Trade Commission and the European Commission. He also served as a Special Assistant Attorney General for the State of Washington in connection with privacy litigation. Reidenberg has chaired the Section on Defamation and Privacy of the Association of American Law Schools (the academic society for American law professors) and is a former chair of the association's Section on Law and Computers.

Prior to coming to Fordham, Reidenberg practiced law in Washington, DC with the international telecommunications group of the firm Debevoise & Plimpton.

Professor Reidenberg received an A.B. degree from Dartmouth College, a J.D. from Columbia University, and both a D.E.A. droit international économique and a Ph.D in law from the Université de Paris -Sorbonne. He is admitted to the Bars of New York and the District of Columbia.