



EDTech Roundtable

**Discussions on two topics:
De-Identification and EDEch apps that
require Parental Consent**



Presenters



Jim Siegl

Future of Privacy
Forum

Why Talk About De-Identification

A VISUAL GUIDE TO PRACTICAL DATA DE-IDENTIFICATION



What do scientists, regulators and lawyers mean when they talk about de-identification? How does anonymous data differ from pseudonymous or de-identified information? Data identifiability is not binary. Data lies on a spectrum with multiple shades of identifiability.



This is a primer on how to distinguish different categories of data.

DEGREES OF IDENTIFIABILITY

Information containing direct and indirect identifiers.

PSEUDONYMOUS DATA

Information from which direct identifiers have been eliminated or transformed, but indirect identifiers remain intact.

DE-IDENTIFIED DATA

Direct and known indirect identifiers have been removed or manipulated to break the linkage to real world identities.

ANONYMOUS DATA

Direct and indirect identifiers have been removed or manipulated together with mathematical and technical guarantees to prevent re-identification.



DIRECT IDENTIFIERS
Data that identifies a person without additional information or by linking to information in the public domain (e.g., name, SSN)



INDIRECT IDENTIFIERS
Data that identifies an individual indirectly. Helps connect pieces of information until an individual can be singled out (e.g., DOB, gender)



SAFEGUARDS and CONTROLS
Technical, organizational and legal controls preventing employees, researchers or other third parties from re-identifying individuals

EXPLICITLY PERSONAL	POTENTIALLY IDENTIFIABLE	NOT READILY IDENTIFIABLE	KEY CODED	PSEUDONYMOUS	PROTECTED PSEUDONYMOUS	DE-IDENTIFIED	PROTECTED DE-IDENTIFIED	ANONYMOUS	AGGREGATED ANONYMOUS
INTACT	PARTIALLY MASKED	PARTIALLY MASKED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
INTACT	INTACT	INTACT	INTACT	INTACT	INTACT	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED	ELIMINATED or TRANSFORMED
NOT RELEVANT due to nature of data	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	CONTROLS IN PLACE	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	LIMITED or NONE IN PLACE	CONTROLS IN PLACE	NOT RELEVANT due to nature of data	NOT RELEVANT due to high degree of data aggregation

SELECTED EXAMPLES

Name, address, phone number, SSN, government-issued ID (e.g., Jane Smith, 123 Main Street, 555-555-5555)

Unique device ID, license plate, medical record number, cookie, IP address (e.g., MAC address 68:AB:6D:35:6D:03)

Same as Potentially identifiable except data are also protected by safeguards and controls (e.g., hashed MAC addresses & legal representations)

Clinical or research datasets where only curator retains key (e.g., Jane Smith, diabetes, Hgb 15.1 g/dL - Cmk123)

Unique, artificial pseudonyms replace direct identifiers (e.g., HIPAA Limited Datasets, John Doe = S1T1 LK619Z) (unique sequence not used anywhere else)

Same as Pseudonymous, except data are also protected by safeguards and controls

Data are suppressed, generalized, perturbed, swapped, etc. (e.g., GPA: 3.2 = 3.0-3.5, gender: female = gender: male)

Same as De-identified, except data are also protected by safeguards and controls

For example, noise is calibrated to a data set to hide whether an individual is present or not (differential privacy)

Very highly aggregated data (e.g., statistical data, census data, or population data that 52.6% of Washington, DC residents are women)

Real-Time Data Shows Exactly How Students Use AI on School Technology

Roughly one-third of student interactions with generative artificial intelligence on school technology involved cheating, self-harm, bullying, and other problematic behaviors, according to data collected and analyzed by Securly, a company offering internet filtering and other safety services.

What's more, Securly identified roughly 1 in 50 student-AI interactions as red flags that students might be involved in violence, cyberbullying, or self-harm.

Securly's analysis looked at nearly 1.2 million interactions in more than 1,300 districts from Dec. 1, 2025, to Feb. 20, 2026.

<https://www.edweek.org/technology/real-time-data-shows-exactly-how-students-use-ai-on-school-technology/2026/03>

De-Identification in FERPA

34 C.F.R. § 99.31(b)(1) — De-identified records and information

An educational agency or institution, **or a party that has received education records or information from education records under this part**, may release the records or information without the consent required by § 99.30 after the removal of all personally identifiable information **provided that** the educational agency or institution **or other party has made a reasonable determination that a student's identity is not personally identifiable**, whether through single or multiple releases, and taking into account other reasonably available information.

Some Terms

Aggregation

Combining data across multiple individuals into summary-level statistics (e.g., counts, averages, distributions) such that individual records are not directly accessible.

De-identification

A set of processes applied to data to remove or modify personal identifiers (direct and indirect) so that individuals cannot be reasonably identified, either alone or in combination with other available information.

Pseudonymization

A specific de-identification technique in which identifiable fields are replaced with consistent, reversible tokens (pseudonyms), where re-identification remains possible using a separately maintained key or mapping.

De-Identification

34 C.F.R. § 99.31(b)(2) — Coded (de-identified) student-level data for research

An educational agency or institution, or **a party that has received education records or information from education records under this part, may release de-identified student level data from education records for the purpose of education research by attaching a code to each record that may allow the recipient to match information received from the same source, provided that—**

- (i) The educational agency or institution or other party that releases de-identified data does not disclose any information about how it generates and assigns a record code, or that would allow a recipient to identify a student based on a record code;
- (ii) The record code is used for no purpose other than identifying a de-identified record for purposes of education research and cannot be used to ascertain personally identifiable information about a student; and
- (iii) The record code is not based on a student's social security number or other personal information.

NDPA on De-Identification

4.5 De-Identified Data Provider agrees not to attempt to re-identify De-Identified Data without the written direction of the LEA. Deidentified Student Data may be used by the Provider for those **purposes allowed under applicable laws**, for the purposes allowed for the processing of Student Data under this DPA, **as well as** the following purposes: (1) assisting the LEA or other governmental agencies in conducting research and other studies; (2) research, development, and improvement of the Provider's educational sites, Services, or applications, and to demonstrate the effectiveness of the Services; and (3) for adaptive learning purpose and for customized student learning.

Provider's use of De-Identified Data shall survive termination of this DPA or any request by LEA to return or dispose of Student Data. **Except for Subprocessors, Provider agrees not to transfer De-Identified Data to any third party unless the transfer is expressly directed or permitted by the LEA or this DPA.** Such Subprocessors must be subject to equivalent terms of the DPA including this one.

Prior to publishing any document that names the LEA, the Provider shall obtain the LEA's written approval of the manner in which De-Identified Data is presented. **If Provider chooses to create De-Identified Data, its process must comply with either NIST de-identification standards or US Department of Education guidance on de-identification.**

“Traditional” De-Identification

- **Blurring:** Reducing the precision of disclosed data to minimize the certainty of individual identification. For example converting continuous data elements into categorical elements that subsume unique cases.
- **Perturbation:** Making small changes to the data to prevent identification of individuals from unique or rare population groups. For example, swapping data among individual cells to introduce uncertainty.
- **Suppression:** Removing data, for example from a cell or row, to prevent the identification of individuals in small groups or those with unique characteristics. Usually requires suppression of non-sensitive data

Privacy Enhancing Technologies (PETs)

Input Privacy

- **Homomorphic Encryption:** A method enabling encrypted computations to be conducted on encrypted data without needing to decrypt it first. The decrypted results match what would have been obtained by performing the calculations on the original unencrypted data.
- **Trusted Execution Environment (TEE):** Secure virtual computing spaces that enable the execution of code and access to data in an isolated manner, detached from the rest of the system. This isolated processing protects against unauthorized access, ensuring the confidentiality and integrity of sensitive data, also known as a secure enclave.
- **Secure Multiparty Computation:** A technique allowing multiple parties to process their combined data without any party needing to share all its information with the others. This approach minimizes the risk of exposing sensitive information.
- **Federated Learning:** A method that allows multiple parties to train AI models on their data and combine identified patterns into a more accurate "global" model without sharing their data.

Output Privacy

- **Differential Privacy:** A method of data anonymization that relies on the injection of "noise" to protect the identification of sensitive, individual data. Differential privacy is commonly used in large data sets.
- **Synthetic Data:** Artificial data created by data synthesis algorithms that mimic the patterns and statistical properties of real data, including personal data, producing comparable results to those obtained from analyzing the original data set.

The Apps not covered by the NDPA

Let's Talk About...

Vendors that Can't be a School Official

Vendors that are not “just” a School Official

- The “it's complicated” vendors
- The car with two steering wheels

Can They Be a School Official?

- **IS IT** something a school would typically use staff to do
- **IS IT** “Materials that are “maintained by an educational agency or institution **or by a person acting for such an agency or institution,**” and contain **“information directly related to a student.”**
- **HOW** would they actually facilitate the school’s FERPA obligation for the parent to inspect?
- **HOW** does the LEA have direct control? (e.g. Admin Dashboard at class, school or district level)

What Are Common “Consent” Use Cases?

- CTE
- Sometimes Sports, Fundraisers
- Consumer Apps (This often impact if State Student Privacy Laws apply)*
- Things you probably should not be using with Students

**"School service" means a website, mobile application, or online service that (i) is designed and marketed primarily for use in elementary or secondary schools; (ii) is used (a) at the direction of teachers or other employees at elementary or secondary schools or (b) by any school-affiliated entity; and (iii) collects and maintains, uses, or shares student personal information. "School service" does not include a website, mobile application, or online service that is (a) used for the purposes of college and career readiness assessment or (b) designed and marketed for use by individuals or entities generally, even if it is also marketed for use in elementary or secondary schools." (Code of VA)*

FERPA CONSENT

§ 99.30 Under what conditions is prior consent required to disclose information?

(a) The [parent](#) or [eligible student](#) shall provide a **signed and dated written consent** before an [educational agency or institution](#) discloses personally identifiable information from the student's [education records](#), except as provided in [§ 99.31](#).

(b) The written consent must:

- (1) Specify the [records](#) that may be disclosed;
- (2) State the purpose of the [disclosure](#); and
- (3) Identify the [party](#) **or class of parties** to whom the [disclosure](#) may be made.

See US Department of Education's [Identity Authentication Best Practices](#) for ways to verify a parent's identity.

Which are Valid FERPA Consent?

- A. A paper that the parent signs/dates that lists the app(s)/App vendor, how the school will use it and types of data collected/created.
- B. Same as above, but a Google form the parent types in their name, email and “I accept”,
- C. Same as above, but a form with a unique ID sent to the parent’s email of record,
- D. A Printed Form that mentions “approved apps” or links to a website that is updated throughout the school year.
- E. Same as above, but an online form with a block to draw a signature
- F. A form that lists app(s)/App vendor, how the school will use it and types of data collected/created and allows the parent to opt-out.
- G. A form or web page that authenticates the parent using a district credential (e.g. a parent portal) that lists app(s)/App vendor, how the school will use it and types of data collected/created.
- H. A paper form that a concurrently enrolled student parent signs/dates that lists the app(s)/App vendor, how the school will use it and types of data collected/created.
- I. “B” combined with notification (email, phone, text) to the parent confirming their consent (not perfect, but better than the “bad” ones)

Which are Valid FERPA Consent?

- A. A paper that the parent signs/dates that lists the app(s)/App vendor, how the school will use it and types of data collected/created. (Valid)
- B. Same as above, but a Google form the parent types in their name, email and “I accept”, (Wrong)
- C. Same as above, but a form with a unique ID sent to the parent’s email of record (Valid)
- D. A Printed Form that mentions “approved apps” or links to a website that is updated throughout the school year. (Wrong)
- E. Same as above, but an online form with a block to draw a signature (Valid)
- F. A form that lists app(s)/App vendor, how the school will use it and types of data collected/created and allows the parent to opt-out. (Wrong)
- G. A form or web page that authenticates the parent using a district credential (e.g. a parent portal) that lists app(s)/App vendor, how the school will use it and types of data collected/created. (Valid)
- H. A paper form that a concurrently enrolled student parent signs/dates that lists the app(s)/App vendor, how the school will use it and types of data collected/created. (Wrong)
- I. “B” combined with notification (email, phone, text) to the parent confirming their consent (not perfect, but better than the “bad” ones) (Valid)

A Starting Point – PTAC Model Terms of Service

- Definition of “Data”
- Data De Identification
- Marketing and Advertising
- Modification of Terms
- Data Collection and Use
- Data Mining
- Data Sharing
- Data Transfer/Destruction
- Data Rights/License
- Access
- Security Controls

https://studentprivacy.ed.gov/sites/default/files/resource_document/file/TOS_Guidance_Mar2016.pdf

Discussion on Next Steps, if any