

Frequently Asked Questions: IDEA Early Childhood - Disclosure Avoidance

Overview

The U.S. Department of Education (Department) established two national technical assistance centers, the Privacy Technical Assistance Center (PTAC) and the Center for IDEA Early Childhood Data Systems (DaSy), to assist States in complying with the privacy, security, and confidentiality requirements of the Family Educational Rights and Privacy Act (FERPA) and the Individuals with Disabilities Education Act (IDEA) as they implement longitudinal data systems. DaSy collaborated with PTAC to adapt and develop resources to help IDEA Part C early intervention and Part B 619 preschool special education staffs address data confidentiality, data sharing, and data privacy questions and issues as they build and use early childhood data systems. This document is part of a series of documents that address data systems issues to specifically meet the needs of IDEA Part C early intervention and Part B 619 preschool special education. Some of the documents in this series were originally developed by PTAC and were adapted by DaSy for this audience.

Purpose

This document is an adaptation of the 2012 release of “Frequently Asked Questions – Disclosure Avoidance” intended for K–12 audiences. It is intended to provide guidance to IDEA Part C early intervention and Part B 619 preschool special education agencies and programs¹ on effective strategies for protecting Personally Identifiable Information (PII) in education and early intervention records^{2,3} in aggregate reports. Presented here in the form of responses to frequently asked questions (FAQs) are suggestions for ensuring that confidentiality requirements are met, particularly compliance with FERPA and IDEA. A list of additional resources follows.

The data in the early childhood system may well enjoy confidentiality under other statutes including the Health Insurance Portability and Accountability Act (HIPAA) of 1996. This guidance is intended to provide general best practices recommendations in this area, rather than guidance specific to HIPAA or other statutes that the Department does not implement.

This document is designed to highlight key issues concerning the use of disclosure avoidance methods. The Department has released additional guidance on best practices for data disclosure avoidance that cover specific strategies in greater depth.

FAQs: Disclosure Avoidance of Personally Identifiable Information in Aggregate Reporting

Question: What are the definitions of *disclosure* and *disclosure avoidance*?

Answer: *Disclosure* means to permit access to or the release, transfer, or other communication of PII contained in any education or early intervention record by any means including oral, written, or electronic, to any party except the party identified as the party that provided or created the record. Disclosure can be authorized, such as when a parent gives written consent to share education or early intervention records with an authorized party (e.g., a researcher). Disclosure can also be unauthorized

¹ Throughout this document, *agencies and programs* refers to state and local entities that administer and/or provide early intervention or preschool special education services under IDEA.

² Under IDEA Part C regulation §303.414(b)(2), an “early intervention record” is considered an “education record” under FERPA. For clarity, we use *education and early intervention records* throughout this document.

³ The U.S. Department of Education’s Family Policy Compliance Office (FPCO) 2012 “Letter to Edmunds” clarifies that the term “early intervention records” in the 2011 IDEA Part C regulations is the same operationally as “education records” under FERPA.

or accidental. An unauthorized disclosure can happen because of a data breach or a loss (see PTAC's "Data Security: Top Threats to Data Protection" brief at <http://ptac.ed.gov/sites/default/files/issue-brief-threats-to-your-data.pdf> for more information and security tips). An accidental disclosure can occur when data released in public aggregate reports are unintentionally presented in a manner that allows individual children⁴ to be identified.

Disclosure avoidance refers to efforts made to reduce the risk of disclosure, such as applying statistical methods to protect PII in aggregate data tables. These safeguards, often referred to as disclosure avoidance methods, can take many forms (e.g., data suppression meaning that some data may not be reported or included in data tables, rounding of data, recoding of data, etc.).

Question: If I am publishing only aggregate data tables, do I still need to be concerned about disclosure avoidance?

Answer: Yes. The aggregation of child-level data into program or service provider level (or state or Federal) reports removes much of the risk of disclosure since no direct identifiers (such as a name, Social Security Number, or child ID) are present in the aggregated tables. Some risk of disclosure does remain, however, in circumstances where one or more children possess a unique or uncommon characteristic (or a combination of characteristics) that would allow them to be identified in the data table (this commonly occurs with small ethnic subgroup populations), or where some easily observable characteristic corresponds to an unrelated category in the data table (e.g., if an early intervention provider or a preschool special education program reports that 100% of males did not improve functioning in one outcome area). In these cases, some level of disclosure avoidance is necessary to prevent disclosure in the aggregate data table.

Question: What legal obligation do IDEA educational and early intervention service programs and providers (including agencies) have to protect PII in aggregate reports?

Answer: Under FERPA and IDEA, education and early intervention service programs and providers (including agencies) reporting or releasing data from education and early intervention records are responsible for protecting PII from disclosure. Additional statutes may add further confidentiality protections. IDEA regulations [§300.642(a) and §303.722(a) for Part B and Part C, respectively] indicate that the data described in section 618(a) of IDEA and its corresponding regulations must be publicly reported by each state in a manner that does not result in disclosure of data identifiable to individual children. Further, because states are required to follow FERPA for Part B and Part C, all FERPA requirements apply.

Question: What issues should IDEA educational and early intervention service programs and providers (including agencies) consider to successfully balance privacy protection requirements with data disclosure requirements?

Answer: Since the release of any data carries at least some element of risk, entirely eliminating the risk of accidental data disclosure may not be possible. However, organizations disclosing the data in the form of public aggregate reports are responsible for minimizing any such risk while still meeting the disclosure requirements and providing as much useful and transparent information to the public as possible. Before each planned release of child data, an organization must determine the acceptable level of risk of disclosure. This means that in each specific case, the entity disclosing the data should evaluate the risk of PII disclosure within the context that the data will be used, and choose the safeguard strategy most appropriate for that particular context.

⁴ Under IDEA Part C regulation §303.414(b)(2)(vii), a "child" receiving services under IDEA Part C is consistent with the definition of a "student" under FERPA, and IDEA Part B regulation §300.8 refers to "a child with a disability." Therefore, we use *child* or *children* throughout this document for consistency with IDEA and more accurate representation of the span from birth through age 5.

Question: Is public reporting of data for small groups (small cells) the same as a disclosure?

Answer: Reporting unrounded frequency counts in small cells, such as an exact number of children in a small group, does not itself constitute a disclosure; however, the smaller the cell size, the greater the likelihood that someone might be able to identify an individual child within that cell and thus the greater the risk of disclosure. Many statisticians consider a cell size of 3 to be the absolute minimum needed to prevent disclosure, although larger minimums (e.g., 5 or 10) may be used to further mitigate disclosure risk.

Question: What standard is used to evaluate disclosure risk?

Answer: The FERPA standard for de-identification assesses whether a reasonable person in the school or IDEA early intervention community who does not have personal knowledge of the relevant circumstances could identify individual children with reasonable certainty based on reasonably available information, including other public information released by an agency, such as a report presenting detailed data in tables with small size cells [34 CFR §99.3 and §99.31(b)(1)]. IDEA state and local agencies and programs should use the “reasonable person” standard to determine whether statistical information or records have been sufficiently redacted before release. This standard is that a reasonable person (i.e., a hypothetical rational, prudent, average individual) in the school or early intervention community should not be able to identify a child because of some well-publicized event, communications, or other similar factor. School and early intervention officials, including teachers, other service providers, administrators, and volunteers, are not considered in making the reasonable person determination because they are presumed to have inside knowledge of the relevant circumstances and of the identity of the children.

Question: What are some commonly used disclosure avoidance techniques?

Answer: Some of the most commonly used disclosure avoidance methods include data suppression, blurring, and perturbation. When deciding which method to apply in a specific situation, it is important to evaluate the different methods in terms of their effects on the utility of the data and the risk of disclosure.

- *Suppression* is removing data (e.g., from a cell or a row in a table) to prevent the identification of individuals in small groups or those with unique characteristics. Suppression may result in very little data being produced for small populations, and it usually requires additional suppression of non-sensitive data to ensure adequate protection of PII (e.g., complementary suppression of one or more non-sensitive cells in a table so that the values of the suppressed cells may not be calculated by subtracting the reported values from the row and column totals). Correct application of suppression generally results in low risk of disclosure; however, it can be difficult to perform properly because of the necessary calculations (especially for large multidimensional tables). Further, if additional information related to the suppressed data is available elsewhere, the suppressed cells may potentially be recalculated.
- *Blurring* is used to reduce the precision of the disclosed data to minimize the certainty of identification. Examples of blurring are rounding, aggregating across different populations or geographies, and reporting percentages and ranges instead of exact counts. Blurring may affect the utility of the data by reducing users’ ability to make inferences about small changes in the data. Similarly, blurring methods that rely on aggregation across geographies or subgroups may interfere with time-series or cross-sectional data analysis. Applying this technique generally ensures low risk of disclosure; however, if any unblurred cell counts or row and/or column totals are published (or are available elsewhere), it may be possible to calculate the values of sensitive cells.
- *Perturbation* involves making small changes to the data to prevent identification of individuals from unique or rare population groups. Examples include swapping data among individual cells (this still preserves the marginal distributions, such as row totals) and introducing “noise,” or errors (e.g., by randomly reclassifying values of a categorical variable). Perturbation helps minimize the loss of data utility relative to other methods (e.g., compared with the complete loss of information with suppression); however, it also reduces the transparency and credibility of the data. Therefore, perturbation is often considered inappropriate for public reporting of program data from a transparency perspective. Applying this technique generally ensures low risk of disclosure as long as the rules used to alter the data (e.g., the swapping rate) are protected. This requires securing the

information about the technique itself as well as restricting access to the original data so that perturbation rules cannot be reverse-engineered.

Question: Does the U.S. Department of Education require IDEA educational and early intervention service programs and providers (including agencies) to use specific data disclosure avoidance techniques?

Answer: The Department does not mandate a particular method, nor does it establish a particular threshold for what constitutes sufficient disclosure avoidance. These decisions are left up to the individual agencies and programs at the state and local level to determine what works best within their specific contexts.

In general, whenever possible, data about children in aggregate publically available reports (e.g., outcomes data presented as cross-tabulated tables) should be combined with data from a sufficient number of other children to disguise the attributes of a single individual. When this is not possible, data on small numbers of children should not be published.

Question: What practical suggestions does the U.S. Department of Education provide to IDEA educational and early intervention service programs and providers (including agencies) to help them implement recommended disclosure avoidance techniques?

Answer: The Department strongly suggests using a computer program to apply disclosure limitation methods because some techniques may be difficult to implement accurately by hand. In particular, to ensure correct application of data suppression, care should be taken when suppressing any complementary cells. Further, it is preferable, from a data user perspective, to apply consistent methods year to year and to use the same disclosure avoidance strategies for similar types of data releases.

Question: Does the U.S. Department of Education provide additional guidance on data disclosure avoidance techniques?

Answer: Yes, the Department has developed best practice guidance for States to consider when designing and adopting their own disclosure avoidance strategies. The “Best Practices for Access Controls and Disclosure Avoidance Techniques Webinar” reviews different disclosure avoidance techniques and their applicability across different contexts, and “Data De-Identification: An Overview of Basic Terms” provides useful definitions of relevant terminology. (See below for direct links to these resources.)

Additional Resources

The Department established the Privacy Technical Assistance Center (PTAC) as a “one-stop” resource for education stakeholders to learn about data privacy, confidentiality, and security practices related to student-level longitudinal data systems. PTAC provides timely information and updated guidance on privacy, confidentiality, and security practices through a variety of resources, including training materials and opportunities to receive direct assistance with privacy, security, and confidentiality of longitudinal data systems. More PTAC information is available on their website: <http://ptac.ed.gov>.

The Center for IDEA Early Childhood Data Systems (DaSy) is a national technical assistance center funded by the Department’s [Office of Special Education Programs \(OSEP\)](#). DaSy works with States to support IDEA Part C early intervention and Part B 619 preschool special education State programs in the development or enhancement of integrated early childhood longitudinal data systems. DaSy’s work is organized around three areas to support IDEA Part C early intervention and Part B 619 preschool special education State staff: (1) generating new knowledge and useful products for States to use in the development and enhancement of statewide early childhood data systems; (2) designing and implementing a continuum of technical assistance strategies with States that are evidence based, relevant, useful, and cost-effective; and (3) providing national leadership and coordination on early childhood data systems. More information about DaSy is available on their website: <http://dasycenter.org>.

Please direct questions to the PTAC at <mailto:PrivacyTA@ed.gov> or 855-249-3072 and/or DaSy at <mailto:dasycenter@sri.com> or 650-859-3881.

The resources below include links to Federal regulations and several guidance and best practices resources. These include some draft recommendations developed by the National Center for Education Statistics (NCES) in published Technical Briefs. While these recommendations may not be appropriate for every situation, they may provide a better understanding of the issues involved in selecting and applying disclosure avoidance methods to education data.

- “Best Practices for Access Controls and Disclosure Avoidance Techniques Webinar.” Privacy Technical Assistance Center (Nov 2012): http://ptac.ed.gov/sites/default/files/Webinar_DD_Nov2012Final.pdf
- “Case Study #5: Minimizing Access to PII: Best Practices for Access Controls and Disclosure Avoidance Techniques.” Privacy Technical Assistance Center (Oct 2012): <http://ptac.ed.gov/sites/default/files/case-study5-minimizing-PII-access.pdf>
- *Code of Federal Regulations - Title 34: Education. Disaggregation of data.* 34 CFR §200.7: www.gpo.gov/fdsys/pkg/CFR-2011-title34-vol1/pdf/CFR-2011-title34-vol1-sec200-7.pdf
- “Data De-Identification: An Overview of Basic Terms.” Privacy Technical Assistance Center (updated May 2013): http://ptac.ed.gov/sites/default/files/data_deidentification_terms.pdf
- FERPA regulations, U.S. Department of Education: www.ed.gov/policy/gen/reg/ferpa
- FERPA regulations amendment. U.S. Department of Education (December 9, 2008): www.ed.gov/legislation/FedRegister/finrule/2008-4/120908a.pdf
- FERPA regulations amendment. U.S. Department of Education (December 2, 2011): www.gpo.gov/fdsys/pkg/FR-2011-12-02/pdf/2011-30683.pdf
- “Frequently Asked Questions—Disclosure Avoidance.” Privacy Technical Assistance Center (Oct 2012): http://ptac.ed.gov/sites/default/files/FAQs_disclosure_avoidance.pdf
- Guidance on the Amendments to the Family Educational Rights and Privacy Act by the Uninterrupted Scholars Act: <http://www2.ed.gov/policy/gen/guid/fpco/ferpa/uninterrupted-scholars-act-guidance.pdf>
- IDEA Part B regulations. U.S. Department of Education (2006). <http://www.ecfr.gov/cgi-bin/text-idx?SID=d74c644d5aeea44a16267317b21601be&node=34:2.1.1.1.1&rgn=div5>
- IDEA Part C regulations. U.S. Department of Education (May 2011): <http://www.ecfr.gov/cgi-bin/text-idx?SID=d74c644d5aeea44a16267317b21601be&node=34:2.1.1.1.2&rgn=div5>
- “Letter to Edmunds (December 2012).” U.S. Department of Education’s Office of Special Education Programs response regarding whether or not “early intervention records” under IDEA Part C are considered “education records” under FERPA. <http://www2.ed.gov/policy/speced/guid/idea/memosdcltrs/edmunds.pdf>
- *SLDS Technical Brief 3: Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting* (NCES 2011-603): <http://nces.ed.gov/pubs2011/2011603.pdf>
- *Statistical Policy Working Paper 22 - Report on Statistical Disclosure Limitation Methodology.* Federal Committee on Statistical Methodology, Office of Management and Budget (1994): <http://fcs.m.gov/working-papers/wp22.html>
- Technical Brief: Statistical Methods for Protecting Personally Identifiable Information in the Disclosure of Graduation Rates of First-Time, Full-Time Degree- or Certificate-Seeking Undergraduate Students by 2-Year Degree-Granting Institutions of Higher Education (NCES 2012-151): <http://nces.ed.gov/pubs2012/2012151.pdf>

Glossary

The 2014 IDEA Part C regulations clarified the relationship between terms used under IDEA Part C and terms used under FERPA. Below is the translation of terms as clarified under IDEA Part C regulation 34 §303.414(b)(2).

For a side-by-side comparison of the primary legal provisions and definitions in IDEA Part B, IDEA Part C and FERPA that relate to the requirement to protect the confidentiality of personally identifiable information of students and children served under the IDEA, please see the IDEA and FERPA Confidentiality Provisions crosswalk available: <http://www2.ed.gov/policy/gen/guid/ptac/pdf/idea-ferpa.pdf>

Crosswalk of Terms

FERPA	IDEA Part C
education record	early intervention record
education	early intervention
educational agency or institution	participating agency
school official	qualified early intervention service (EIS) personnel/service coordinator
State educational authority	lead agency
student	child under IDEA Part C

FERPA Definitions

- **Education program** is defined as any program principally engaged in the provision of education, including, but not limited to, early childhood education, elementary and secondary education, postsecondary education, special education, job training, career and technical education, and adult education, and any program that is administered by an educational agency or institution. For more information, see the Family Educational Rights and Privacy Act regulations, [34 CFR §99.3](#).
- **Education records** means records directly related to a student and maintained by an educational agency or institution, or by a party acting on behalf of the agency or institution. For more information, see the Family Educational Rights and Privacy Act regulations, [34 CFR §99.3](#).
- **Personally identifiable information (PII)** from education records includes information, such as a student's name or identification number, that can be used to distinguish or trace an individual's identity either directly or indirectly through linkages with other information. See Family Educational Rights and Privacy Act regulations, [34 CFR §99.3](#), for a complete definition of PII specific to education records and for examples of other data elements that are defined to constitute PII.
- **Early childhood education program** means- (a) A Head Start program or an Early Head Start program carried out under the Head Start Act ([42 U.S.C. 9831 et seq.](#)), including a migrant or seasonal Head Start program, an Indian Head Start program, or a Head Start program or an Early Head Start program that also receives State funding; (b) A State licensed or regulated child care program; or (c) A program that—(1) Serves children from birth through age six that addresses the children's cognitive (including language, early literacy, and early mathematics), social, emotional, and physical development; and (2) Is—(i) A State prekindergarten program; (ii) A program authorized under section 619 or Part C of the Individuals with Disabilities Education Act; or (iii) A program operated by a local educational agency. For more information, see the Family Educational Rights and Privacy Act regulations, [34 CFR §99.3](#).
- **Educational agency or institution** means any public or private agency or institution to which this part applies under [§99.1\(a\)](#). (Authority: 20 U.S.C. 1232g(a)(3)). For more information, see the Family Educational Rights and Privacy Act regulations, [34 CFR §99.3](#).

IDEA Part B and Part C Definitions

- **Child**, as defined by Part C regulations, means an individual under the age of six and may include an infant or toddler with a disability, as that term is defined in [34 CFR §303.21](#). For more information, see the Individual with Disabilities Education Act regulations, [34 CFR §303.6](#).
- **Child with a disability**, as defined by Part B regulations, means a child having mental retardation, a hearing impairment (including deafness), a speech or language impairment, a visual impairment (including blindness), a serious emotional disturbance (referred to in this part as “emotional disturbance”), an orthopedic impairment, autism, traumatic brain injury, another health impairment, a specific learning disability, deaf-blindness, or multiple disabilities, and who, by reason thereof, needs special education and related services. For more information, see the Individual with Disabilities Education Act regulations, [34 CFR §300.8](#).
- **Education records**, as defined by Part B regulations, mean the type of records covered under the definition of “education records” in 34 CFR part 99 (the regulations implementing the Family Educational Rights and Privacy Act of 1974, 20 U.S.C. 1232g [FERPA]). ([34 CFR §300.611](#))
- **Early intervention records**, as defined by Part C regulations, mean all records regarding a child that are required to be collected, maintained, or used under Part C of the Act [IDEA] and the regulations in this part. ([34 CFR §303.403](#))
- **Participating agency**, as defined by Part B regulations, means any agency or institution that collects, maintains, or uses personally identifiable information, or from which information is obtained, under Part B of the Act [IDEA]. ([34 CFR §300.611](#))
- **Participating agency**, as defined by Part C regulations, means any individual, agency, entity, or institution that collects, maintains, or uses personally identifiable information to implement the requirements in Part C of the Act [IDEA] and the regulations in this part with respect to a particular child. A participating agency includes the lead agency and EIS [early intervention service] providers and any individual or entity that provides any Part C services (including service coordination, evaluations and assessments, and other Part C services), but does not include primary referral sources, or public agencies (such as the State Medicaid or CHIP [Children's Health Insurance Program]) or private entities (such as private insurance companies) that act solely as funding sources for Part C services. ([34 CFR §303.403](#))
- **Personally identifiable**, as defined by Part B regulations, means information that contains: (a) the name of the child, the child's parent, or other family member; (b) the address of the child; (c) a personal identifier, such as the child's social security number or student number; or (d) a list of personal characteristics or other information that would make it possible to identify the child with reasonable certainty. ([34 CFR §300.32](#))
- **Personally identifiable information**, as defined by Part C regulations, means personally identifiable information as defined in [34 CFR 99.3](#) [See FERPA], as amended, except that the term “student” in the definition of personally identifiable information in [34 CFR §99.3](#) means “child” as used in this part and any reference to “school” means “EIS [early intervention service] provider” as used in this part. ([34 CFR §303.29](#))

The contents of this document were developed under a grant from the U.S. Department of Education, #H373Z120002 and in collaboration with the following offices and centers:

- At the U.S. Department of Education:
 - Office of Special Education and Rehabilitative Services (OSERS), Office of Special Education Programs (OSEP)
 - OSERS' Office of Policy and Planning (OPP)
 - Office of the General Counsel (OGC)
 - Office of Management (OM), Privacy, Information, and Records Management Services (PIRMS)
 - OM's Family Policy Compliance Office (FPCO)
- The Center for IDEA Early Childhood Data Systems (DaSy)
- The Privacy Technical Assistance Center (PTAC)

Although many offices within the U.S. Department of Education provided input into, and review of, the content in this document to make available technical assistance on best practices, the content should not be read as representing the policy of, or endorsement by, the U.S. Department of Education. For further information, you may contact the DaSy grant project officers, Meredith Miceli and Richelle Davis.