To: MLDSC Staff

From: Dawn O'Croinin, Counsel

Date: March 3, 2015

Re: Dashboard and Reporting Suppression Rules

As we move forward towards publishing new dashboards in support of addressing the MLDSC policy questions, and ultimately publishing research papers, it is important for all agency staff to understand the legal obligations of the MLDSC with respect to publishing information containing small cell sizes. It is also important for staff to understand the role of the Office of the Attorney General within this process.

The Office of the Attorney General is headed by Attorney General Brian Frosh, who represents Maryland as the chief legal officer of the State of Maryland. The agency contains several divisions within the central office, along with specialized units, designed to represent and protect the citizens of Maryland. Those divisions include, for example: antitrust, civil litigation, consumer protection, contract litigation, criminal appeals and educational affairs. The OAG also provides counsel to agencies throughout the State. Although I am assigned to the MLDSC, and I am considered "authorized staff" pursuant to Md. Code, Educ. Art., §24-703(d), I work for the OAG. My role is to provide legal advice to the agency regarding all operational matters, including labor and employment, data privacy, the impact of proposed legislation, contracts and procurement, the open meetings act, and the public information act. My goal is to prevent the agency from involvement in litigation, or from sanctions by the U.S. Department of Education under FERPA, 20 U.S.C. §1232g, or the Department of Labor under 20 C.F.R. Part 603.

The primary statutes of concern when determining what information may be displayed within a dashboard or research report are Md. Code, Educ. Art. §24-701, *et seq.* (the MLDSC authorizing statute); the Family Educational Rights Privacy Act (FERPA), 20 U.S.C. §1232g, and its accompanying regulations, 34 C.F.R. §99.1, *et seq.*; the Federal State Unemployment Compensation Program, 20 C.F.R. Part 603; Md. Code, Labor & Employment Art. §8-625; and Md. Code, General Provisions Art. §4-101, *et seq.* (Maryland Public Information Act). These statutes must be read and analyzed together in order to provide appropriate guidance regarding data suppression.

The MLDSC statute provides that the "Center may only use de-identified data in the analysis, research, and reporting conducted by the Center." Md. Code, Educ. Art., §24-703(g)(2). The data used in the release of reports and in response to data requests must also be aggregate. Md. Code, Educ. Art., §24-703(g)(3). The Center may not release information that may not be disclosed under FERPA or any other relevant privacy laws and policies. Md. Code, Educ. Art., §24-703(g)(5). And, the statute specifically provides, "Data that may be identifiable based on the size or uniqueness of the population under consideration may not be reported in any form by the Center." The restrictions on data use and

reporting outlined above apply equally to workforce and education data under the MLDSC authorizing statute.

With respect to FERPA, the activities of the MLDSC fall within the "audit or evaluation" exception. *See* 34 C.F.R. §99.35; *see also* Letter from Dale King, Director, Family Policy Compliance Office at the United States Department of Education Office of Management to Ms. Dawna McIntyre, Associate Counsel for University of Massachusetts and Catherine Schultz, Principal Counsel for the Maryland Higher Education Commission (November 22, 2013).

The United States 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." This definition is used consistently throughout the federal government.   It does not matter whether the "other personal or identifying information" about a specific individual is public or non-public. The ultimate goal is to prevent a disclosure of an individual's identity when using de-identified, aggregate data to answer the research and policy questions of the MLDSC in conformity with the statutory mission of the agency, *see* Md. Code, Educ. Art. §24-703(f), under the "audit and evaluation" exception of FERPA.

I have reviewed the National Center for Education Statistics' SLDS Technical Brief entitled, "Statistical Methods for Protecting Personally Identifiable Information in Aggregate Reporting" (December 2010, Brief 3). I have also reviewed the Maryland ESEA Flexibility Accountability Addendum (September 19, 2013) provided to me by the Maryland State Department of Education. Both documents provide guidance for data suppression in statistical reporting. The standard for Maryland, according to the ESEA Accountability Addendum is that "Maryland does not report the results of any subgroup smaller than 10 in number." (Addendum at 17). Where the value is less than 10, an asterisk is used and an appropriate footnote is provided. *Id.* As noted in the NCES brief, as of winter 2010, a survey of all states and the District of Columbia showed that 39 states use a minimum reporting group size of 10 students. *See* NCES Brief at 7.[1] Based on these two documents, and my reading of FERPA and the MLDSC authorizing statute, I recommend that the MLDSC use a minimum reporting group size of 10 for all research and dashboards.

The minimum reporting number, however, is not the only effort the MLDSC must make to protect the privacy of student data in reporting results. The NCES brief advises that "[t]he minimum number for categories within subgroups of reporting can be set lower than the size of the subgroup minimum, but there should be a minimum size specified for individual categories to guard against unintentional disclosures." *See* NCES Brief at 7. This may present problems in relation to other data

---

[1] Seven states set the minimum reporting group at five, and the remaining five states use a higher value, ranging from 15 to 30.

reported, even where a minimum size for the category within a subgroup is set at 5 (as one of the suggestions in the Brief). My recommendation is to keep the minimum reporting group size at 10, including for categories within subgroups.

With respect to reporting percentages, the NCES Brief recommends that an entity "[u]se only whole numbers when reporting the percentage of students for each category of an outcome measure (e.g., the percentage assessed)." This increases the difficulty of reverse calculating the data to obtain identifying information. The NCES Brief further recommends recoding categories with percentages expressed to protect student privacy as follows:

a. **For reporting variables/outcome measures with more than 300 students and no related subgroup with fewer than 200 students, use the following approach:**
    i.   Recode categories with values of 99 to 100 percent to greater than or equal to 99 percent ($\geq$ 99 percent).
    ii.   Recode categories with values of 0 to 1 percent to less than or equal to 1 percent ($\leq$ 1 percent).
    iii.   Otherwise, report the percentage of students in each category using whole numbers.
b. **For reporting variables/outcome measures with 201 to 300 students and no related subgroup with fewer than 200 students, use the following approach:**
    i.   Recode categories with values of 98 to 100 percent to greater than or equal to 98 percent ($\geq$ 98 percent).
    ii.   Recode categories with values of 0 to 2 percent to less than or equal to 2 percent ($\leq$ 2 percent).
    iii.   Otherwise, report the percentage of students in each category using whole numbers.
c. **For reporting variables/outcome measures in which the number of students ranges from 101 to 200**, use the following option in this group and all related subgroups with more than 200 students:
    i.   Recode categories with values of 98 to 100 percent to greater than or equal to 98 percent ($\geq$ 98 percent).
    ii.   Recode categories with values of 0 to 2 percent to less than or equal to 2 percent ($\leq$ 2 percent).
    iii.   Recode the percentage in each remaining category in all reporting groups or subgroups to intervals as follows (3-4, 5-9, 10-14, 15-19, …, 85-89, 90-94, 95-97).
d. **For reporting variables/outcome measures in which the number of students in the smallest reporting group or subgroup ranges from 41 to 100**, use the following option in that group or subgroup and use option c for each related reporting group or subgroup with more than 100 students:
    i.   Recode categories with values of 95 to 100 percent to greater than or equal to 95 percent ($\geq$ 95 percent).
    ii.   Recode categories with values of 0 to 5 percent to less than or equal to 5 percent ($\leq$ 5 percent).
    iii.   Recode the percentage in each remaining category in all reporting groups or subgroups to intervals as follows (6-9, 10-14, 15-19, 20-24,…85-89, 90-94).

e.   For reporting variables/outcome measures in which the number of students in the smallest reporting group or subgroup ranges from 21 to 40, use the following option for that group or subgroup, use option d for each related reporting group or subgroup with 41 to 100 students, and use option c for those with more than 100 students:

   i.   Recode categories with values of 90 to 100 percent to greater than or equal to 90 percent (> 90 percent).

   ii.   Recode categories with values of 0 to 10 percent to less than or equal to 10 percent (< 10 percent).

   iii.   Recode the percentage in each remaining category in all reporting groups or subgroups to intervals as follows (11-19, 20-29, …, 80-89).

f.   For reporting variables with 10 to 20 students in the smallest subgroup, use the following option for that group or subgroup, use option e for each related group or subgroup with 21 to 40 students, use option d for those with 41 to 100 students, and use option c for those with more than 100 students:

   i.   Collapse all outcome measures to only two categories, using the same collapsing rules across all subgroups for each outcome measure (e.g., assessment results collapsed to below the proficient level and at or above the proficient level by sex, racial and ethnic groups, disability status, etc.).

   ii.   Recode categories with values of 0 to 20 percent to less than or equal to 20 percent (< 20 percent), and recode the other category to greater than 80 percent (> 80 percent).

   iii.   If both collapsed categories have percents of 21 to 79 percent, recode the percentage in each collapsed category to intervals as follows (21-29, 30-39, …, 70-79).

NCES Brief at 28-29. For purposes of the research and dashboards of the MLDSC, I do not advise recoding as a matter of course. Although we may recode and follow this protocol for certain types of information reporting (for example, assessments with breakdown by race, gender, etc.), it is not necessary at all times. Instead, I recommend that the MLDSC not report percentages which yield a number of less than 10 out of the population studied. This will keep the reporting consistent with the minimum reporting number.

At all times the data reported should be check by reverse calculations to determine whether it is possible to identify a small cell size of unreported data based on the other data which is, in fact, reported. This will satisfy the MLDSC obligations under both FERPA and the MLDSC authorizing statute. And, it is important to note that there may be circumstances where the numerical value of a subgroup is greater than or equal to 10 but must be suppressed anyway in order to prevent a disclosure. As noted in the NCES Brief, "the goal is to publish summary results that do not allow someone to learn information about a specific student." NCES Brief at 4. This is helpful in conceptualizing the legal obligations of the Center without becoming fixated on the phrase "personally identifiable information" as the sole limiting factor.

This memorandum has focused on the MLDSC obligations under FERPA and the MLDSC authorizing statute. With respect to the labor and employment data, the Center's obligations are similar, except that the Center must prevent the disclosure of the identity of the individual as well as specific employers. This topic will be explored in greater detail as the Center's workforce data expands and we begin to see research results that will allow me to better frame the discussion.

If anyone should have any questions, please do not hesitate to contact me. My direct office line is (410) 706-1092, and my e-mail address is: dawn.o'croinin@maryland.gov.