

Jan Kiehne, CT Board of Regents for Higher Education

Jan_kiehne@hotmail.com

P20 WIN Program Manager & Lead Tester

Kumar Murugappan, CT Board of Regents for Higher Education

kmurrugappan@commnet.edu

P20 WIN Developer and Tester

Benjamin Klimczak, CT Board of Regents for Higher Education

bklimczak@commnet.edu

CC IRDB Lead Analyst

Teresa Boyd-Cowles, CT Board of Regents for Higher Education

Boydt@ct.edu

CSU IR Repository Lead Analyst

Laurel Buchanan, University of Connecticut Health Center

lbuchanan@uchc.edu

P20 WIN Host System Administrator

Ajit Gopalakrishnan, CT State Department of Education

Ajit.Gopalakrishnan@ct.gov

SDE Representative to P20 WIN Data Governance Board and analyst

Chuck Martie, CT State Department of Education

charles.martie@ct.gov

SDE representative to P20 WIN Data Steward Group and analyst

SECTION 3 –ALIGNMENT OF PURPOSE& SCOPE

Provide a brief summary of the data request:

This project will test the validity of the P20 matching algorithm by comparing data matched through P20 WIN against a separately matched set of results between high schools, state universities and community colleges in order to evaluate enrollment in their programs.

Clearly state the purpose of this request and describe how the purpose is an audit or evaluation of federal or state supported education program(s) (See Audit or Evaluation Exception 20 U.S.C. 1232g(b)(1)(C), (b)(3), and (b)(5) and §§99.31(a)(3) and 99.35)

This data request has the simple purpose of evaluating the enrollment of Connecticut State University and Community College education programs.

Describe the benefit that this audit/evaluation will provide to the state or to a specific agency which is an education related member of the Board.

As a test evaluation, the benefit to the state is primarily in assessing P20 WIN system functionality and identifying areas in need of improvement. In order to maximize the investment being made in the system, the Participating Agencies and stakeholders need to trust the data matching process and the quality of the data output. Conducting this evaluation will provide an understanding of system functionality necessary for conducting future audits and evaluations.

Provide a brief description of the method for analysis.

Descriptive statistics will be used to determine variance by school, institution, gender and race/ethnicity.

Processes for manipulating the data will include:

- Actively linking data tables together to validate the presence or absence of surrogate keys. (Note: this is done without the presence of primary personal identifiers such as name or SSN. Rather, this is accomplished primarily through secondary or surrogate keys assigned by the system and not tied to source data.)
- Filtering and sorting data to understand how rows in data sets are unique or duplicated.
- Modifying data configuration to establish ideal parameters for future analysis. (e.g. date ranges, field values, sequencing)
- Matching data aggregated from the output to known aggregated data from existing repositories to provide input into data quality

Provide a description of the documents or reports that will reference data from this query. For each report indicate the audience for the report and expected indicators or measures to be included in each.

There will be no public report utilizing student data from this data query. Rather, information obtained about the quality of the data matches will be used by the P20 WIN participating agencies to inform future data requests for additional audits or evaluations.

In addition to utilizing a minimum cell size as documented by the Data Governance Procedure, identify/describe the statistical methods that will be used to minimize the risk of re-identification of PII for data to be published. Some typical methods are listed below; however, this list is not comprehensive. Options should be used in combination for maximum security. See the [USED Technical Brief 3](#) as a resource.

- **Suppression by:** cell, row, sampling (present data for a portion of students: e.g. 80%)
- **Blurring through:** aggregation of groups, rounding, use of ranges, use of top/bottom categories (e.g. 'less than 5%, greater than 95%),
- **Perturbation:** data swapping, adding noise, use of synthetic data

n/a – This data request will not be used to produce any public reports utilizing student data.

SECTION 4 – PROTECTION OF CONFIDENTIALITY

By what date will the resultant dataset and all paper or electronic copies will be destroyed by the requestor?

Policy requires that the resultant data set be destroyed within 12 months from the date the dataset was downloaded. Note, that the Data Governing Board may set a shorter or longer time frame for destruction of the dataset. This 12 month limitation may be extended with written approval from the Data Governing Board. Requestors need to submit the Query Management Document as a 'continuation' for approval of an extension.

All resultant data will be destroyed by **7/31/3014**.

With the addition of your initials by each item, indicate that you agree to the following:

RJM
RJM
RJM

I will only utilize the data received through this query to meet the purpose as described. The approval given to receive data through this query does not confer approval to use it for another purpose.

I will not re-disclose the data received through this query approval process to any public official who has not been authorized by the Data Governing Board to receive it, and who has not also signed a Personal Statement of Confidentiality and Non-Disclosure.

I will take the necessary and appropriate precautions to safeguard personal information and will comply with all state and federal laws concerning the safeguarding and disclosure of such information.

BQH
BQH
BQH
BQH
BQH
BQH
BQH

I will not use this dataset to re-identify individuals.

I agree that approval to receive data through this query does not convey ownership of the data.

I agree that prior to the public release of any documents or reports generated from this query I will supply all reports or documents to the Data Governing Board for review and verification that the intended purpose has been adhered to.

I agree to store all query data, in print or electronic form, in a locked receptacle that can be accessed by authorized persons only.

I agree to store all query data on secure desktop computers and in secure files to which access is restricted to authorized persons only.

I agree that no query data may be transmitted via email or placed or stored on a mobile computing or storage device. For purposes of this agreement, a definition of "mobile computing device" includes, but is not limited to, notebooks, palmtops, PDAs, iPods®, Blackberry® devices, and cell phones with internet browsing capability. A "mobile storage device" includes but is not limited to, mobile computing devices, diskettes, magnetic tapes, external/removable hard drives, flash cards (e.g., SD, Compact Flash), thumb drives (USB keys), jump drives, compact disks, digital video disks, etc.

BQH

I agree to destroy all confidential information obtained through this agreement as soon as such information is no longer needed.

BQH

I agree that the Data Governing Board will be allowed access to monitor all authorized users to ensure such users observe the confidentiality requirements of the information obtained under this Agreement.

BQH

I agree to comply with all provisions of the P20 Win Data Query Management Procedure.

SECTION 5—ADDITIONAL INFORMATION

Provide any additional comments that would be useful to the Data Governing Board in considering this request.

Since this is a test data request and test evaluation, it is expected that the query will may need to be repeated multiple times before establishing that the system can run without error. BOR is requesting the ability to run the same test evaluation query up to 15 times as a part of the testing process.

SECTION 6 – SIGNATURE OF REQUESTOR AND DATE

12.1 Signature of Principal Requestor

I understand that the entities that are providing data to P20 WIN (hereinafter Participating Agencies) have made reasonable efforts to ensure that the data available through P20 WIN are up-to-date, accurate, complete and comprehensive at the time of disclosure. These records reflect data as reported to the Participating Agencies by their data-submitting organizations for the reporting period indicated. Changes or updates to the data may occur after the time of disclosure and may impact data that have previously been made available. The Participating Agencies are not responsible for data that are misinterpreted or altered in any way. Derived conclusions and analyses generated from this data are not to be considered attributable to the Participating Agencies or the participating organization(s) from which the data originated.

I certify that the information supplied in this form, with attachments, is complete, accurate. The analysis will be conducted according to the protocol approved by the Data Governing Board, applicable federal, state and local laws regarding the protection of education records and unemployment insurance records. I will ensure that all protocol changes will be prospectively reviewed by the Data Governing Board. I will request approval from the Data Governing Board for changes to the study and will not implement proposed changes until I receive Data Governing Board approval. I will promptly report to the Data Governing Board any related complaints, problems, and/or breaches of confidentiality.

Signature B. J. Hesch Date 6/25/13

CT Preschool through Twenty & Workforce Information Network (P20 WIN)

DATA GOVERNING BOARD REVIEW DOCUMENT - **QUERY REQUEST**

This form and all attachments are completed by the P20 WIN Data Governing Board in response to a query request.

SECTION 1 – BASIC REQUEST INFORMATION

Application Submission Date:

TYPE OF SUBMISSION: Initial Application Continuation Repeat Modification

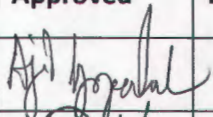
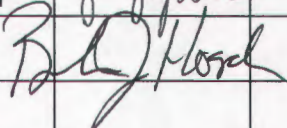
P20 WIN Query Number: P20W_1305_2_0002.n

Query Title: SDE-BOR Test Evaluation – A

Agencies from which data is being requested (Check all that apply):

SDE BOR DOL

The Policy Lead from Each Participating Agency Completes indicates their approval or denial of the query to enable an audit/evaluation of education programs according to FERPA.

Agency	Name	Approved	Denied	Comment	Date
SDE	Ajit Gopalakrishnan				6/20/13
BOR	Braden Hosch				6/20/13

Is each agency in agreement? Yes No

Next steps:

- If the participating agencies whose data would be included in this specific data request each approve the query, then representatives from each of the included agencies must sign a Memorandum of Agreement for a P20 WIN Data Request.
- Data Governing Board Review Document is shared with the requestor
- Data query process is aborted or continues based upon the decision made.