



## Rhode Island Longitudinal Data System Data Governance Committee

### DRAFT Minutes

Wednesday, April 17, 2024 – 10:00 AM

Zoom Meeting: <https://uri-edu.zoom.us/j/98015003108>

This meeting was RECORDED

RI Department of Education, Room 501  
255 Westminster St, Providence, RI, 02903

---

#### COMMITTEE MEMBERS IN ATTENDANCE:

Dana **Brandt** (RILDS); Rebecca **Bucci** (OHHS Delegate); Jessica **Cigna** (SOS); Vincent **Flood** (DOA); Scott **Gausland** (RIDE); Harris **Hameed** (DOIT); Mike **Matkowski** (OMB Delegate); Andrea **Spargo** (RIOPC); and Megan **Swindal** (DLT).

#### COMMITTEE MEMBERS ABSENT:

Colleen **Caron** (DCYF).

#### GUESTS:

None.

#### SUPPORT STAFF:

Kim **Pierson** (RILDS).

**Dana Brandt** called the April 17, 2024, meeting of the Rhode Island Longitudinal Data System Data Governance Committee to order at 10:02 AM EST, acknowledging that a quorum was present.

1. **WELCOME** - **Brandt** opened the meeting and welcomed everyone.
2. **VOTE** on Adoption of Meeting Minutes – **Brandt** tabled the approval of the March RILDS Data Governance Committee meeting minutes until the May Committee meeting.
3. **DISCUSSION** of & **VOTE** on Data Quality & Metadata Policies – **Brandt** introduced the next order of business, the Data Quality & Metadata Policies.

**Brandt** reviewed the Data Contributing Agency responsibilities described in the Data Quality & Metadata Policies.

The first policy describes Data Collection & Preparation. **Brandt** explained that the RILDS Center is asking the Data Contributing Agencies to identify at least one Data Steward for each data

collection shared through RILDS. The Data Steward is responsible for communicating a schedule for updating their data collection(s), including both update frequency and deadlines, to the RILDS Center. The Data Steward is also responsible for providing definitions and metadata for each collection, including documentation of any calculations, deduplications, formatting changes, and other processing the Steward performs before transferring data. These documentation requirements are not intended to add to Data Steward burdens, but to allow for seamless transition if/when Data Stewards or file formats change.

**Swindal** asked for clarification that the data update schedule would be set by the Data Steward or Data Contributing Agency. **Brandt** confirmed that the Data Contributing Agency and/or Data Steward would determine the schedule.

**Hameed** inquired about the process of updating definitions when definitions change for recurring data collections. **Gausland** suggested that Data Stewards could maintain the timelines where certain definitions are applicable.

**Cigna** asked about the preferred format for Data Stewards to get the metadata to the RILDS Center engineers. **Brandt** replied that Data Stewards would receive the file from the RILDS Center in Excel or CSV format so they could edit that file and return it to the Center.

**Brandt** reviewed the second policy which covers Data Quality. She highlighted that within 15 days of receiving data extracts, the RILDS Center performs data quality checks. The three checks performed include ensuring that (1) no unexpected elements were received, (2) all expected elements were provided, and (3) the record or row counts are consistent with previous extracts for a comparable period.

In instances where the Center receives unexpected data elements, they are treated as unauthorized data, and the Center will immediately destroy the file received and request a new data extract from the Data Contributing Agency to be received within fifteen (15) business days of notification.

In situations where expected fields are missing the Center notifies the Data Steward of the missing elements and requests a corrected extract. The Data Steward is responsible for transferring a corrected extract within fifteen (15) business days of notification.

When the Center discovers the record or row counts are inconsistent with previous extracts, the Center contacts the Data Steward to confirm whether the extract is complete (e.g., were records mistakenly excluded). If the Data Steward confirms that the extract is complete, the Center asks the Data Steward to document the programmatic or policy changes that affected the record or row counts (e.g., middle schools are now required to report grades earned where previously only high schools reported grades earned). If the Data Steward confirms that the extract is missing records, the Center proceeds with import and requests a corrected extract to be transferred within fifteen (15) business days of notification.

**Spargo** inquired if these checks would begin before or after the metadata is received from each Data Steward. **Brandt** confirmed that until the metadata is received full checks cannot be currently completed.

**Hameed** inquired if the RILDS Center would, upon receipt of unexpected elements, return the extract to Data Contributing Agencies. **Brandt** stated that only the field names would be retained, and the extract would be destroyed.

**Swindal** asked how data quality issues outside of the three stated checks would be handled if encountered during analysis or use. **Brandt** replied that it depended on how critical the data quality issue is, but that she would like to create a policy around critical issues.

**Gausland** asked how the Center documents data quality issues that they discover. Do they create additional metadata he inquired. **Brandt** explained that the Center's internal data dictionary contains fields for documenting notes about fields and data collections.

**Brandt** reviewed the third policy, Data Modification. She explained that if the file layout changes, the Center will follow the process outlined in their MOU with the individual Data Contributing Agency to modify the data elements shared. In most agreements, the procedure is for the Center to write a letter to the Point of Contact listed in the Data Contributing Agency MOU specifying the new field(s) or collection(s) to be added. The Data Contributing Agency's Point of Contact signs and returns this memorandum, authorizing the modification. Brandt noted that the Center retains all records and elements, until or unless the Data Contributing Agency requests destruction.

**Bucci** inquired about how to handle fields in the data table layout that get depreciated and removed. **Brandt** stated that the RILDS Center would be notified that the field has been removed and it would be reflected in the Data Modification memorandum procedures previously discussed, but the field name and data would be preserved in the RILDS Center's data table and model. She elaborated that the field name would not need to be manually added to each file's extract by the Data Contributing Agency because the Center's engineering team would note the field to be expected missing on import.

**Brandt** requested a motion to approve sending the Data Quality & Metadata Policies to the RILDS Executive Governing Committee for review and approval subject to the amendments proposed today. **Gausland** motioned, **Spargo** seconded, and the Committee unanimously approved to send the Data Quality & Metadata Policies minutes to the Executive Governing Committee for review and adoption.

**VOTING IN FAVOR:** Brandt, Bucci, Cigna, Flood, Gausland, Hameed, Matkowski, Spargo, Swindal

**VOTING IN OPPOSITION:** None

**ABSTENTION:** None

**ABSENT:** Caron

#### 4. **DISCUSSION** of & **VOTE** on Data Requests

- a. *van Bueren Charitable Foundation, Greater Newport Dashboard*

**Brandt** introduced the van Bueren Charitable Foundation (vBCF) request for a dashboard

following students from selected municipalities who graduate from Rhode Island public schools through postsecondary, workforce training, and employment.

**Swindal** flagged the “CIP SOC” issue regarding the research question “How many and what percentage of these graduates enter employment in an industry aligned to their field of study?”. The CIP SOC Crosswalk matches postsecondary programs of study to occupations requiring those skills or knowledge to be successful. She explained that it's not always clear in the data as someone might graduate in a particular field and then go into teaching within that field, but the teaching industry is not likely aligned to reflect their field of study. **Brandt** stated that the Center would report sector using NAICS codes given that DLT employment and wage records do not contain SOC codes and given the timeline for the project.

**Hameed** flagged issues related to limited reporting of gender data in some datasets. **Brandt** explained that this cohort that is being followed will reflect the demographic data as reported at the time of high school graduation and is a required data element in public K-12 education data.

**Cigna** asked about what type of disclaimers are usually included on dashboards produced by the Center and how the impact of COVID disruptions would be noted in this Output. **Brandt** responded that, unless the requestor asks otherwise, the Center is not flattening demographic fields so that the data can more fully reflect on an individual's identity at any given point in time.

**Gausland** and **Cigna** flagged that it was a large request with a short timeline. **Brandt** noted that the Center would be breaking the information release into phases with some information requested by May and some in June. She also mentioned that the Center is also looking into hiring a contractor to support this work given their other pending requests. **Brandt** suggested that in a future meeting the Committee establish a prioritization matrix within the many outstanding requests based on the value of the information not just to the requestor, but to the State.

**Hameed** asked if the Center had all the data needed to complete the request. **Brandt** stated that the RILDS Center has all the data needed except for updated federal PIRL (Participant Individual Record Layout) data from DLT which they are in the process of modifying their MOU to receive. **Swindal** flagged that the completion criteria for PIRL may make use of that data less than ideal for this request and a follow up conversation should occur between DLT and the Center.

**Gausland** asked if the Center would have to set aside other committed work to meet the van Bueren Charitable Foundation deadline. **Brandt** confirmed that the FY2023 SLDS grant work would not begin until the vBCF work was completed. She highlighted that the Request and Release Policy prioritized state agency requests over third researcher party requests, but that the RILDS Executive Governing Committee suggested prioritizing requests based on funding.

**Gausland** acknowledged that this type of request showcases the work that the RILDS is built to perform.

**Swindal** asked **Brandt** to make a note to follow up with vBCF to clarify if they are requesting apprenticeship in general or Registered Apprenticeship, a specific program.



**Cigna** asked if the scripts created by the vBCF would be proprietary or could they be used for other analyses, possibly looking at these questions statewide which may justify the tight timeline and bumping down the list of other important projects. **Gausland** agreed that this information processed statewide would be very helpful. **Brandt** stated that she didn't know where the IP law and state syntax paid for with private foundation funds falls, but she believed the Center owned their work.

**Swindal** suggested having a conversation with the Data Steward in the DLT workforce training data collection area.

The Data Contributing Agencies unanimously approved.

**VOTING IN FAVOR:** Cigna, Gausland, Spargo, Swindal

**VOTING IN OPPOSITION:** None

**ABSTENTION:** None

**ABSENT:** Caron

## 5. ADJOURNMENT

There being no further business, the April 17, 2024, RILDS Data Governance Committee meeting adjourned at 10:42 AM EST.