



**Solicitation Information**

**May 9, 2014**

**Addendum #2**

**RFP # 7548642**

**TITLE: HIGHER EDUCATION DATA WAREHOUSE**

**Submission Deadline: May 23, 2014 @ 2:00 PM (Eastern Time)**

**PLEASE NOTE THAT THE SUBMISSION DEADLINE HAS BEEN EXTENDED TO FRIDAY, MAY 23, 2014 AT 2:00 PM (ET).**

**VENDOR QUESTIONS WITH STATE RESPONSES ARE ATTACHED. NO FURTHER QUESTIONS WILL BE ANSWERED.**

**Gail Walsh  
State of Rhode Island  
Division of Purchases**

## **Vendor A**

1. Are full proposals publicly posted? Could elements of the proposal (for example, not-yet-launched work samples) be excluded/redacted from public posting?

A: Yes- if posted, pieces could be redacted.

## **Vendor B**

### **Page 4 Background/Overview**

1. Other than the 2006 effort to build a data warehouse, have there been other efforts to build a more robust warehouse that have not succeeded?

A: No, but there has been additional work to link and report the 3 separate systems.

See: <http://ridatahub.org/datamart/indicator-report/exploring-rhode-islands-public-college-students/>

2. What was the budget and expenditure on the original effort in 2006?

A: \$200,000 per year to update and maintain a limited data set

3. What specific planning/budget decisions are expected to be made from this data warehouse? What data is desired that the current system does not provide today (for example, "Cost per student credit hour")?

A: That will be determined by the Technical Advisory Group (TAG)

4. Only Operating budgets are mentioned in this background section. Is there a need or consideration for analysis of Capital budgeting or is this out of scope?

A: At this stage, this component is out of scope.

5. Are the data requests driving this procurement available to the vendor for review as amendments to the RFP to ensure a complete proposal response?

A: There are no official requests available. However, higher education often receives questions regarding remediation, retention, dual enrollment, academic outcomes for students who receive scholarships and financial aid, etc.

6. There is mention of approximately 700K biographic/term records, implying a level of granularity no more detailed than a student registered in a term. Does the level of detail of a student registered in a given section/course exist today?

A: Yes.

7. Will the current data warehouse remain in service or is the expectation the result of this project will entirely replace the current one?

A: There is currently a database which has served as a warehouse for specific reporting purposes. This will be replaced.

- a. If replaced, what expectation is there for the data to be incorporated into the new data warehouse? Is there a conversion expected?

A: The data will be in the new data warehouse and conversion will be necessary.

### **Page 4 Scope of Work/Tasks**

8. The first paragraph refers to "campus-wide data warehouse reporting and extract solutions". Today does each college have its own data warehouse and then this is aggregated into one state wide data warehouse? Is this the architecture you would require moving forward?

A: This is the current architecture and would be the architecture going forward.

9. Is the Rhode Island Department of Education data warehouse going to continue and still need to be fed or is it in an end of life cycle to be replaced by office of higher education warehouse?

A: They are separate data warehouses and continue to be so.

### Page 5 Scope of Work/Tasks

10. Can you clarify what is meant by “automatically compile live data from multiple platforms”? Should the data warehouse query live data in a federated manner or is a latency of an overnight extract-transform-load (ETL) process acceptable?  
A: TBD by TAG, but at a minimum monthly
11. What proportion of resources for the project can come from campus IT departments to support design, implementation, and reporting requirements?  
A: Extremely limited. A Technical Advisory Group (TAG) will be able to meet according to a set schedule.

### Page 5 Project Schedule

12. In regards to the work plan for the first 6 months. Does the OHE have an expectation of what would be delivered at a minimum? Additionally, is the OHE open to a rapid development lifecycle (for example, Agile) for delivering iterative components?  
A: OHE is interested in proposals that suggest an accelerated approach.

### Page 5 Deliverables

13. Item #1 refers to the ability for the data warehouse to “automatically sync data across all systems into a centralized location using Student ID number...”
- a. Does RI have a shared student ID number that travels with the student across all attendance at any public RI HE institution?  
A: No.
  - b. If not, how does RI accomplish record matching and avoid duplicates in the current HE data warehouse?  
A: Matching process
  - c. If the logic is present, would it be available for us to review?  
A: After award
14. Item #2 refers to a technical advisory group.
- d. Does this group exist today?  
A: In AdHoc form only
  - e. Is this the same as the HE technical steering group and the HE program advisory committee mentioned in the Purpose on page 4 – or are these 3 distinct groups?  
A: Technical Advisory and Technical Steering are the same. In addition there will be a program advisory committee created to inform the project from the program side. A steering committee oversees the entire longitudinal data system (LDS) grant and includes members across state agencies.
  - f. How does RI envision the coordination between an external vendor and these internal groups? For example, will there be an internal project manager / executive sponsor to help manage?  
A: There may be funds for an internal project manager. However, the vendor should propose the method by which project management would be performed by their entity.
15. Item #2. What is meant by “updates data across all systems according to set rules...”?  
A: Updates should be systematic across institutions based on rules established by the TAG.
- g. Is there an expectation of write-back from the data warehouse to the source systems based on outcomes of reporting/analysis/predictive model scoring?  
A: No.

- h. Can you elaborate on what is meant by federal and state data calendars and how they would affect the update schedule for the data warehouse?

A: There are many collections and most are available on the web. Below are links to Complete College America and IPEDS:

<http://completecollege.org/>

<https://surveys.nces.ed.gov/ipeds/>

16. Item #6. Can you expand on what is meant by “sets up a connection to an analytics program that will allow for real time data analysis for those not using the raw data”?

A: The build needs to be extensible so that a front end can be added on in the future.

- i. Are you just looking for the ability to connect SAS or SPSS to the data warehouse models or is this something more?

A: Also CSV, SSRS and excel

- j. Additionally, real-time is an important term here. Do you expect the data warehouse information to be an up to the minute reflection of your source systems?

A: At least monthly updates will be required.

- k. What degree of latency will RI be comfortable with for the data warehouse?

A: At least monthly updates will be required.

#### Additional Questions

17. Is there an existing Data Governance group?

A: No, there is a cross-agency steering committee.

18. Will the state provide access to the vendor to source requirements docs used to develop this RFP?

A: Yes.

- a. For example, data warehouse metadata, the ER Diagrams/data models/dictionaries for the purposes of constructing the best and most comprehensive technical proposal to meet the needs?

A: Yes.

19. What are the source systems at all institutions? Are they the same? Ellucian Banner? PeopleSoft?

A: Oracle/PeopleSoft at Rhode Island College and the University of Rhode Island. Banner at the Community College of Rhode Island. All systems are hosted locally at each institution.

PeopleSoft Banner

20. Are there existing technologies in place and available to be used to construct the data warehouse or is the full \$600K max budget for ALL software and services?

A: Yes

- b. What are the existing technologies?

- c. Database Solution / Platform

A: MSSQL 2008R2

- d. Extract, Transformation and Load Tools

A: **SQL Server Integration Services (SSIS)** Microsoft **SQL Server** database software that can be used to perform a broad range of data migration tasks.

- e. Single Sign-on Solution

A: Customized single sign on application, built by Intrafinity. The core is built on SharePoint, Active Directory, and SAML 2.0 compliant applications.

- f. Reporting Tools / Solution  
A: **SQL Server Reporting Services (SSRS)** is a server-based report generation software system from [Microsoft](#). SSRS enables users to quickly and easily generate reports from [Microsoft SQL Server](#) databases.
21. Is the hardware and infrastructure already available and procured?  
A: Yes.
- g. What is the hardware platform and specifications?  
A: There is adequate server capacity, a sufficient operating system and database platform. There is not adequate back-up. We expect that the vendor propose the specifications necessary for the build.
    - i. # of servers, CPUs, Memory
    - ii. Physical vs. virtual
    - iii. Operating System
    - iv. Database Platform
22. If the hardware and infrastructure are not already available and procured, will the vendor be responsible for this piece as well and is this included in the \$600K budget?  
A: No.
- h. If the vendor would be responsible for this are there dependencies on other technologies or preferences that the vendor should be aware of?  
A: Please see page 3, Q 20.
23. How many end users are expected for phase one? In the Office of Higher Ed? The campuses?  
A: TBD
24. Does this new data warehouse replace any existing campus data warehouses?  
A: No.
25. What is the survey and list of primary/secondary data sources? In our experience there are typically numerous shadow systems on campuses and they need to be reviewed and addressed or else there is an incomplete picture of data. This is especially true if an existing data warehouse or reporting database has not been meeting requirements for a number of years.  
A: A data dictionary exists.

**Place of Performance**

26. Does the purchasing entity favor or require local / onsite presence for any or all of this project? If so, what proportion of time?  
A: Onsite management is estimated at 50% FTE

**Proposal Submission**

27. Are there page limits on the supplemental information? For example, the key personnel CV's, are they included in the 10-20 pages of the technical proposal or just their overview and the key CVs can be attached as supplemental?  
A: There are not page limits but we recommend extreme caution in choosing applicable and helpful information.

**Vendor C**

1. What are the Student Information Systems at each of the institutions?  
A: Please see Q 19 on p.3.
2. What is the format/frequency of institution data files?  
A: Currently term; going forward at a minimum monthly
3. How is Accuplacer data delivered to RIOHE? Periodic flat files?  
A: Flat files

4. Should SAS/SPSS have the ability to source data directly from the Data Warehouse?  
A: Yes.
5. What is the ETL Tool and Database platform?  
A: Please see Q 20 on p.3.
6. Can we get a copy of the OHE's data model, and please describe the subject areas required for this phase of the project. For example are these in scope? Any sample reports to support this is appreciated as well.  
A: The following link contains CEDS elements which is the foundation of the data model.  
<https://ceds.ed.gov/elements.aspx> . All of the items below are in scope.
  - a. Student Enrollment
  - b. Student Applicant
  - c. Assessment Score
  - d. Student Term records
  - e. Course Catalog
7. From a staffing perspective, do the resources have to be onsite and/or offshore or can some of the team be leveraged offshore?  
A: We expect on-site management, but off shore work is acceptable.

## **Vendor D**

1. What data warehousing tools (RDBMS, ETL, BI) are currently in use with the existing data warehouse?  
A: The current data warehouse offered no tools. Reports will built by the vendor on a set schedule.
2. Does RIOHE expect the vendor to enhance the existing data warehouse, replace it or build a new and separate data warehouse?  
A: Build new with structure developed by ProvPlan – see  
<http://ridatahub.org/datamart/indicator-report/exploring-rhode-islands-public-college-students/>
3. Does the budget of \$600,000 include any required software (RDBMS, ETL, BI) or hardware?  
A: No. Software/ hardware are either currently available or will be purchased as necessary by the state.
4. What vendor assisted RIOHE in the development of the current data warehouse?  
A: Center for Labor Markets & Policy, Drexel University
5. What are the significant data categories or data elements that are missing from the existing data warehouse?  
A: This will be determined by the TAG. At a, minimum, all CEDS and IPEDS elements will be included. The build should be flexible to add components in the future.
6. What is the role of RIDE in the project?  
A: RIDE is the fiscal agent and serves on the steering committee.
7. Can RIOHE share the data model that has been created in ER/Studio Data Architect? How complete is this data model?

A: This is attached as a pdf. ER Studio model is about 50% complete. Currently, tables related to organization, person, class and admission are listed. Tables pertaining to IPEDS, faculty, programs and academics still need to be created.

8. The information on Page 10 (Elements of Proposal) includes “5. Fee estimate, expenses to be incurred and payment schedule”. Please confirm that this information should be in the separate cost proposal and not included in the technical proposal. Please confirm that the other 6 items in this section should form the outline for the technical proposal.

A: Please provide in a separate cost proposal.

9. Page 9 of the RFP includes the Technical Proposal/Required Elements list. Can RIOHE please provide clarification on item 3 “Capacity of the Agency Effectively to Administer the Project”?

A: What is your experience with similar project of scope and size? What is your ability to manage technical groups and work cross- agency?

Assumptions:

1. RIOHE will be responsible for all hardware and software including RDBMS, ETL and BI software.

A: Yes.

2. RIOHE will be responsible for developing extract files from the RIOHE source systems in the agreed upon format.

A: That is the responsibility of the vendor.

3. There will be data from 3 primary sources (University of Rhode Island, Rhode Island College and Community College of Rhode Island).

A: Yes.

4. Historical data will be limited to the current academic year.

A: No.

5. RIOHE will provide work space for ...team.

A: Yes, for the selected vendor.

6. The proposal does not include the development of any reports or reporting semantic layers.

A: Only SSRS

7. RIOHE will provide access to the required individuals for the development of requirements.

A: Yes.

8. The project will begin approximately June 1, 2014 and conclude by November 30, 2014.

A: Conclude September 30, 2015

9. RIOHE will complete their tasks according to the agreed upon project schedule.

A: Yes.

10. All source system extracts will include SSN which will be utilized to match records.

A: No.

11. The ...team will work on-site at RIOHE approximately 50% of the time.

A: Yes.

12. RIOHE will provide remote access to the data warehouse environment.

A: Yes.

13. The estimated number of ETL jobs and complexity is listed below:

A: TBD

- a. Simple – 10 (Loading dimensions)  
A: TBD
- b. Medium – 10  
A: TBD
- c. Complex – 10  
A: TBD

## **Vendor E**

1. Regarding Data Integrity— do we need to provide the capability to Profile and correct information from the source systems? For example:
  - a. Missing data items that need to be remediated (ie SSN)?  
A: Data audits are in scope.
  - b. Duplicate data elements that need to be matched and merged (ie student id's, course id's)?  
A: Data audits are in scope.
  - c. Data items that need to be validated (i.e. address)?  
A: Data audits are in scope.
2. Would the ongoing data reporting needs be better served through real time (vs ETL batch) updates to the data warehouse?  
A: Either is acceptable but this will be determined by the Technical Advisory Group.
3. Please identify all the Source and Target data sources the system will integrate to/from (ie Oracle, MS SQL Server, Spreadsheets).  
A: Please Qs 19 and 20 on p.3.
4. Please identify the hardware and operating system that is intended to be used for the system.  
A: Please see Q20 on p.3.
5. The RFP references a requirement for 'firm fixed price' and also asks for estimated hours and hourly rates for services to be provided. Will the billing/payment mechanism for services be; fixed price (i.e. milestone or deliverable) based, time and materials (i.e. based on hours worked), or time and materials - not to exceed (i.e. capped)?  
A: It is a firm fixed price which must include all of these items.
6. The first task within the services scope of work includes the analysis required to determine the detailed requirements for the full project. Could a vendor propose a fixed price for the initial task / phase (up to and including the detailed requirements having been determined and documented), and subsequently provide pricing for the remaining effort (according to the required contracting model – per question # \_\_\_ above)?  
A:
7. General Questions
  - a. Do you currently use Banner from Ellucian as the primary source of information?  
A: Please see Q19 on p.3.
    - i. If not, what Administrative system is used at each institution?  
A:
    - ii. If so, is the data is in an Oracle database at each source?  
A:
8. Does RIOHE or RIDE currently have an Information Technology department that supports all current applications and databases "in-house"?

- A: RIDE and each of the higher education institutions have a technology department.
- a. If not, what applications are not supported in-house?
 

A: Applications are available in-house but RIDE cannot support them for this project.

    - i. Would data from these applications be included in the new data warehouse?
 

A: Data from the higher education institutions would be included. Data from RIDE would not.
    - ii. Where are the other applications hosted?
 

A: Applications are hosted in-house.
  - b. Will any of this staff be dedicated to this project full-time?
 

A: Limited staffing will be available to coordinate advisory groups and track project status.
9. Is the “Not-to-Exceed” budget for Professional Services relating to building the new Data Warehouse only (and excludes travel, additional software licenses and/or hardware acquisitions)?
- A: Yes.
10. Current Environment (built in 2006)
- a. What is the database type and version of the existing database environment (e.g., Microsoft SQL Server 2005, Oracle 9i)?
 

A: Please see Q20 on p.3.
  - b. How many different environments comprise the current environment (e.g., Development, Test, Quality Assurance, Production)?
 

A: This is not applicable as it will be a new build based on RIDE infrastructure. OHE data warehouse.
  - c. How much data is in each environment (e.g., 500Gb, 2Tb)?
 

A: For the data currently in the database, see: <http://ridatahub.org/datamart/indicator-report/exploring-rhode-islands-public-college-students/> There is no OHE data warehouse.
  - d. What is the historical growth rate of data been for the last three years (e.g., 20% year over year)?
 

A: Consistent
  - e. Source Data:
    - i. What are the application names and versions from which data is sourced?
 

A: Please see Q19 on p.3. Banner and PeopleSoft
    - ii. For each source application
 

A: Banner and PeopleSoft

      1. What are the database types and versions and operating systems and versions?
 

A: Banner: The current Banner version is 8.5.6 soon to be upgraded to 8.6.3.

and PeopleSoft
      2. How many different tables/files serve as source data (e.g., five tables)?
 

A: *student, human resources, and financials.*

3. Approximately how many columns/fields of data are sourced from the application (e.g., 30 columns)?

A: TBD

4. What's the frequency that data is extracted from the application (e.g., hourly, daily, weekly, 10<sup>th</sup> day)?

A: At a minimum monthly.

5. Please indicate whether each application is an RIOHE internal or external data source.

A: The institutions host internally.

iii. ETL

A: Both

1. What technology and methods are used to move data from the sources systems to the current database?

A: Flat file transfer

2. How many business or transformation rules currently exist?

A: TBD

3. What is the level of complexity (high, medium or low)?

A: TBD

iv. Is RIOHE satisfied with the ETL technology and methods currently employed?

A: No.

f. Users

i. How do users access the current environment?

A: They do not. There is a public platform on the data hub. No users have access to the Drexel Database.

ii. Approximately how many users does the current environment support?

A: None

iii. How many concurrent users are there?

A: None

g. Documentation

i. Please describe the documentation that exists for the current environment (e.g., logical architecture, physical architecture, data architecture, business requirements, functional requirements, system requirements, data mapping, logical and physical data models).

A: The ER Studio model is attached. A data dictionary exists on the RI DataHUB.

ii. Is it complete and current?

A: The ER Studio model is approximately 50% completed.

h. Reporting and Analytic Tools

i. What are the names and version of the reporting and analytic tools currently used?

A: None

ii. Is RIOHE satisfied with this technology?

A: No.

i. High Availability / Disaster Recovery

i. Does the current environment have high availability or disaster recovery as part of the infrastructure? If so, please describe that as well as the processes that support it.

A: It is part of the Drexel University infrastructure.

j. Data Quality

i. Please describe the technical data quality processes that are in place.

A: N/A- This is a new build.

ii. Please describe the organizational data quality processes that are in place (e.g., Data Ownership and Stewardship processes, Standards Boards).

A: This is a grant deliverable.

iii. Please describe the relative data quality of the data as it exists in the source systems.

A: Overall data quality is fair; Data have become better in recent years.

iv. Are like data consistently represented across the various source systems (e.g., Student Ids, Names, state codes, course codes, gender codes)?

A: Yes.

v. Please describe any data quality issues that might exist in the current database environment.

A: This will be made available upon grant award.

k. Existing Issues

i. What types of problems are you experiencing with this current database system (i.e., slow report performance, data volumes are limited, incomplete or bad data, hard to maintain, etc.)?

A: The system is used only by the designer who provides reports as specified.

ii. What can't you do that you would like to do with this current database system?

A: The data warehouse should be able to help each institution have a better sense of student flow through the RI higher education system, which should help in enrollment planning, for example.

## 11. New Environment

a. Does RIOHE have a preferred target database type and version that it wishes to implement (e.g., SQL Server 2014, Oracle 12G, Teradata 14.2)?

A: Please see Q20 on p.3.

i. If not, does RIOHE expect the vendor to facilitate an assessment process that compares various database vendors?

A: na

b. How many different environments will comprise the new environment (e.g., Development, Test, Quality Assurance, Production)?

A: A data warehouse build would necessarily include all of these environments.

c. What is the anticipated growth rate of data for the next three years (e.g., 20% year over year)?

A: TBD

d. Source Data:

i. If the new environment will support data that is sourced from applications other than those listed under Current Environment above, please list the application names and versions and database types and versions from which data will be sourced?

A: TBD

ii. For each source application please estimate

1. What are the database types and versions and operating systems and versions?

A: Please see Qs 19 and 20 on p.3.

2. How many different tables/files serve as source data (e.g., five tables)?

A: TBD

3. Approximately how many columns/fields of data are sourced from the application (e.g., 30 columns)?

A: TBD

4. What's the frequency that data is extracted from the application (e.g., hourly, daily, weekly, 10<sup>th</sup> day)?

A: At least monthly.

5. Please indicate whether each application is an RIOHE internal or external data source.

A: Each data warehouse at the colleges are hosted locally.

e. Users

i. Approximately how many users will the new environment support?

A: TBD, but for this phase a small number is anticipated.

ii. How many concurrent users are anticipated to be using the system?

A: Not more than 50.

f. Requirements

i. Have the business requirements for the new solution been defined, documented and agreed-upon?

A: No.

g. High Availability / Disaster Recovery

i. If required, have the high availability and disaster recovery requirements and processes been defined?

A: No.

h. Historical Data Retention

i. Please describe the historical data retention requirements.

A: Data from 2006

i. Security

- i. Please describe security requirements for the new environment.

A: MS active directory

12. Scope of Work

- a. In order to design the system to meet future needs, please describe the anticipated use of the system in follow-on phases.

A: This question is unclear

b. Tasks

- i. In paragraph 1 of Tasks, it says, "Using the RI Department of Education's data warehouse model as a guide, the foundation for the OHE's warehouse has been created in ER/Studio Data Architect." Please the "foundation" in more detail.

A: Please see attached RIDE Data Flow model.

1. What elements from CEDS, IPEDS and institution data files have been incorporated?

A: All are being incorporated but it is incomplete.

2. How far along has the model been developed?

A: It is approximately 50% complete.

3. Is it expected that the vendor will build on this model or will the vendor build entirely new models?

A: Yes, the model is a starting point.

4. Will vendor need to provide its own copy of ER/Studio or can it leverage the licenses owned by RIOHE?

A: It will be provided.

ii. ETL

1. What technology and methods will be used to move data from the sources systems to the current database?

A: That is the vendor's responsibility.

2. How many new business or transformation rules will be built?

A: TBD

3. What is the level of complexity (high, medium or low)?

A: TBD

4. Does RIOHE have a preferred target ETL tool that it wishes to employ as part of the new solution?

A: See Q20 on p.3.

5. Does it prefer to develop all ETL logic as part of the capabilities within the selected target database?

A: Yes

iii. Reporting Requirements

1. How many campus users will need to be consulted regarding daily and long-term reporting requirements and analytics?

A: That will be determined in consultation with the Institutions, but it will be limited.

2. How many report writers and architects will need to be consulted to model the data and structure the ETL's?

A: This is vendor-determined.

3. Is there an existing process in place that vets and approves requested requirements and designs?

A: This process will be developed by the vendor working with the TAG.

iv. Data Warehouse Build

1. Please describe what is meant by "live data" (e.g., sub-second, minute, or hourly access to data that has been added or changed in a source system).

A: At least monthly

13. Work Location

a. Does RIOHE have any expectation or rules regarding work performed remotely or offshore?

A: 50% on-site project management

14. Does RIOHE anticipate that an existing team will participate in the implementation of the solution in any way? If so:

A: There is not an existing team but it is anticipated that the three Institutional Researchers (IRs) will be involved. It is the vendor's responsibility to determine a work and participation schedule.

a. Please describe how the existing team will participate.

A: The vendor will determine participation working with the TAG. The existing team consists of the three IRs. It is the vendor's responsibility to determine a work and participation schedule.

b. Please describe the roles, skills and technologies with which the existing team has experience.

A: Existing team consists of the three IRs. It is the vendor's responsibility to assess experience.

15. Is it a goal of RIOHE that once the solution is built and put in place that an existing team will assume ongoing management and support of the solution?

A: The vendor may propose a maintenance fee, not to be included in the cost proposal, that articulate the cost to maintain after the project is completed.

a. If so, please describe the roles, skills and technologies with which the existing team has experience, if different from above.

A: This team would need to be developed.

## **Vendor F**

1. Have you already selected the target software platform for your Data Warehouse?  
A: Please see Q20 on p.3.
2. If not, have you identified any software selections or standards for the following Data Warehouse Components:
  - a. Database  
A: Please see Q20 on p.3.
  - b. Extract, Transformation and Load Tools  
A: Yes.
  - c. Single Sign-on  
A: Yes.
  - d. Reporting  
A: No.
  - e. Dashboards  
A: No.
3. Have you considered or evaluated using any packaged data marts or analytics with pre-built connections to your existing ERP systems?  
A: Not applicable at this time.
4. Is there a specific reason for the requested timeframe or would an accelerated approach be of interest?  
A: An accelerated approach is welcome.

## **Vendor G**

1. Scope of work - We understand that there is an existing data warehouse. Please share the Current Architecture, Technology / Tool Stack.  
A: Please Q20 on p.3.
  - a. Also share any key challenges faced in the current environment  
A: Limited staffing and time
    - i. Scope of work - Does the scope of work is to build a new data warehouse from scratch or extend / enhance the existing data warehouse?  
A: Build new, with many inherited components that will expedite the build.
    - ii. Scope of work - If the new data warehouse is to be built from scratch, then can you please share the preferred technology stack  
A: Please see Q20 on p.3.
      - Database Platform  
A:
      - ETL Tools  
A:
      - Reporting / BI Tools  
A:
      - Analytical tools (Predictive Model)  
A: N/A
    - iii. Scope of work - Is Vendor expected to do Tool Evaluation exercise and recommend suitable tool set for DB / ETL / BI?  
A: No, only if you believe the current solution is unacceptable- or would like to propose a more viable and sustainable one.

- iv. Scope of work - Does Rhode expect the Vendor to perform Capacity Planning exercise to recommend the sizing for hosting hardware and storage needs?  
A: Yes.
- v. Scope of work - Please provide details about the source system
- # of source systems  
A: 3
  - Subject Areas covered  
A: IPEDS, Complete College America as well as items included in ER Studio model (attached). In addition further elements may be added based on the discovery phase.
  - Approximate # of tables to be considered for data extraction from each system  
A: TBD
- vi. Scope of work - How do you assess the quality of data present in the source system on a scale of 1- 5? ( 1 – Extremely good, 3-Good, 5 – Extremely bad)  
A: Good
- vii. Scope of work - Does the scope include onetime migration of data from the source systems? If Yes, how many years of historical data have to be migrated to the proposed higher education Data Warehouse?  
A: 8 years
- viii. Scope of work - The number of Downstream applications that are expected to extract data from the proposed Data Warehouse  
A: TBD
- ix. Scope of work - Any preference on the Cloud based hosting solutions?  
A: No
- x. Scope of work - Does the reporting environment require internationalization in terms of language, currencies, etc.?  
A: No.
- xi. Scope of work - What is the estimated number of reports to be developed from Data Warehouse? Please classify them in terms of
- A: Not in scope
- Canned Reports  
A: Not in scope
  - Ad-hoc Reports  
A: Not in scope
  - Analytical Reports  
A: Not in scope
- xii. Scope of work - Please provide details on the target user group who will use the reports? Is the report access limited to internal users or open for public domain viewing?  
A: It is limited to internal users who will develop reports for public consumption.
- xiii. Scope of work - Please share data audit, compliance requirements (if any).  
A: To be developed

- xiv. Scope of work - Are there any specific solution requirements (e.g. Department wise reporting to be delivered) that need to be completed in Phase1? Is there any road map available? If yes, please share the same.  
A: This will be negotiated with the vendor upon award of contract.
- xv. Page 4: Background/Overview - Current data volume information is provided in the RFP document. Can you also provide us with the expected YOY data growth?  
A: TBD
- xvi. Page 4: Scope of work – The RFP states on a requirement for “Operational Reporting”, Can you please provide us with the information on the load strategy/frequency. E.g.: daily, hourly, etc.,  
A: At a minimum, monthly.
- xvii. Page 4: Scope of work - Are there any real-time reporting requirements?  
A: No.
- xviii. Page 4: Scope of work - There is a reference of "Web based system" for data consumption. Please confirm if the expectation is to build user friendly using web based application in Java /.NET based technologies or using BI tools like Business Objects, Cognos, and SSRS etc. technologies?  
A: Please see Q20 on p.3. RIDE uses .Net, Csharp, SSRS.
- xix. Page 4: Scope of work - Is production support required after the go-live? If yes please provide the details around the type of support - 24x7/ 16x5/ only business hours.  
A: Business hours through the end of the contract
- xx. Page 5: Project Schedule - Please provide us with the specific timeline within which the project is expected to be completed.  
A: The project must be completed by June 30, 2015
- xxi. Page 5: Project Schedule - Can you provide additional information on Work Plan (WP)? Does the WP include schedule alone?  
A: This will be negotiated.
- xxii. Page 5: Project Schedule - Please provide details on the start date for Phase 1 & Phase 2.  
A: This depends on the vendor’s approach to the build.
- xxiii. General - RFP states "While our long-term plans include aligning an analytics and reporting tool" - Please confirm if a reporting tool is either finalized or tool is in use already?  
A: Neither. A reporting tool is in a subsequent phase.
- xxiv. General - How many stakeholders would be involved based on the subject areas / sources to be incorporated? Please share details on their level of involvement and the availability of the stakeholders & SMEs.  
A: Existing team consists of the three IRs. It is the vendor's responsibility to assess experience.

- xxv. General - Our understanding is that all stakeholders are stationed in one location and the requirement gathering exercise would be performed in that location. Please provide details of locations if the requirements have to be gathered from stakeholders located in multiple locations.  
A: The majority of this work can be conducted from the Office of Higher Education in Providence. However, there may be instances in which requirements gathering would occur at each of the three higher education institutions.
- xxvi. General - Please confirm if there is any preferred mode of project execution: Agile, Waterfall, etc...  
A: No
- xxvii. General - Is there is any preferred location for project execution? (e.g.: specific place in US or Offshore, etc.)  
A: There is an expectation of 50% management on-site.
- xxviii. General - Please share the internal team availability and the expected Roles and Responsibilities of the vendor team.  
A: This will be determined by the vendor in coordination with the Institutional Researchers and the Office of Higher Education.

## **Vendor H**

1. The RFP states there is an existing data collection process in place. Would the proposed solution utilize the existing solution infrastructure for transport and extraction?  
A: The existing solution is likely not viable for this project.
2. The RFP requires building new business rules for the extraction process. What percent of the business rules are in place for the existing data collection solution?  
A: Though some business rules are in place, each business rule will need to be revisited in light of CEDS.
3. The RFP states that there was some foundational work completed on the Data warehouse modeling for the proposed solution. How many data domains and attributes already exist, and can you estimate the percentage of completion for the existing foundation?  
A: Please see attached PDF. It is approximately 50% completed.
4. The RFP does not include requirements for reporting and dashboards for the historical and operational data. Is reporting/dashboards outside of the scope of this RFP?  
A: Yes.
5. The RFP does not include requirements for the hardware and software for the system. Are there any specific technology and platform requirements?  
A: Please see Q20 on p.3.
6. The RFP does not include requirements for hosting the system? Does the vendor need to provide hosting or will this be hosted by the agency?  
A: It will be hosted by the agency.
7. The RFP does not include requirements for configuration management and environment (i.e. development, functional testing, user acceptance testing, and production) requirements. What are the environments that will be setup for the proposed system? Is the vendor responsible to setup the environment and deploy the solution in all the environments?  
A: Yes

8. The RFP does not include requirements for managing user identities. How will the users access the proposed system?  
A: There will be limited users and RIDE does have a user identity management system, which is available as necessary.
9. The RFP does not include requirements for security and encryption. Are there any specific guidelines for the proposed solution?  
A: RIDE uses MS Active Directory. The vendor is encouraged to propose additional recommendations.
10. The RFP does not include requirements for Training and Help. Is this handled by the agency?  
A: There will be a limited amount of training / and a training manual that will be a deliverable of the vendor.

## **Vendor I**

1. RFP: "Data Warehouse with data from source systems"  
Please provide a list of the required source systems for inclusion in the warehouse? We require it to accurately estimate the timeline and budget the project. We would appreciate the following;
  - a. the name of the system  
A: Please see Q19 on p.3.
  - b. the type of data within the system to be brought into the RIOHE Data Warehouse  
A: Please see Q19 on p.3.
  - c. and whether it is locally or vendor hosted  
A: Please see Q19 on p.3.
2. RFP: "Longitudinally (from 2005-2012), there are approximately 700K Biographic and Term records, 2.7M Course records plus 400K Course Catalog records, 19K Major Catalog records, and 300K Accuplacer records."
  - a. Will data from the current system, described above, be a source system for the new system?  
A: Yes
  - b. How much historical data is required for this new data warehouse?  
A: 8 years
3. Please define the State and Federal regulatory reporting requirements within the scope of this RFP.  
A: There are many collections and most are available on the web. Below are links to Complete College America and IPEDS:  
<http://completecollege.org/>  
<https://surveys.nces.ed.gov/ipeds/>
4. Please define the financial analysis to be performed within the scope of this RFP.  
A: TBD
5. Outside of the State and Federal regulatory reporting, and the financial analysis noted above, what are the other purposes or uses of this Data Warehouse?  
A: It is a tool for the Office of Higher Education and each of the three institutions to make improvements in programs and services.

6. RFP: "While this database satisfied a need at the time, new requirements and data requests necessitate a much more robust system".  
Would you please elaborate on the above "new requirements and data requests"?
- A: Any new proposed legislation is accompanied by ad hoc data requests. In addition, initiatives such as Complete College America add to the requirements.
7. RFP: "Using the RI Department of Education's data warehouse model as a guide, the foundation for the OHE's warehouse has been created in ER/Studio Data Architect".  
What do you mean by the "the foundation for the OHE's warehouse has been created..."?
- A: Please see the attached PDF.
8. RFP: "The OHE model incorporates elements from the CEDS postsecondary data model, IPEDS and institution data files".  
Would you please share the list of data elements required in the data warehouse, from CEDS postsecondary, IPEDS or other institutional data files?
- A: Please see the data dictionary available on the RI DataHUB. In addition, you can visit the IPEDS, CEDS and Complete College American sites for more information.
9. RFP: "Consult with campus users to determine daily and long-term reporting and analytics requirements. Consult with data warehouse report writers and architects to determine how to model the data and structure the ETLs to streamline report writing."
- a. In order to accurately scope the proposal and the expected deliverables, how many reports and analytics requirements are expected to be delivered as part of this proposal?
- A: CSV, SSRS, excel, SAS, SPSS
- b. Does the warehouse need to support IPEDS Federal Reporting or do you expect the vendor to produce the actual IPEDS extracts/reports?
- A: The DW should support IPEDS
- c. Or is all reporting and analysis outside the scope of this RFP?
- A: Reporting by the vendor is out of scope.
10. Can you provide any specific details on reporting and analysis business requirements which is within the scope of this RFP?
- A: Not at this time
11. Who are the perceived users of this Data Warehouse?
- A: Internal users for this phase of the project.
12. Are you open to a cloud offering or does it have to be locally hosted?
- A: The project is flexible and would consider a cloud offering, but local hosting is provided in-kind.
13. RFP: "Develops single sign-on procedures so that confidential data is accessible only by authorized users."
- a. What will the Single Sign-on integrate with, is there an existing portal, and so, is Portal Integration expected?
- A: Single sign on can be but does not have to be integrated with the existing portal.
- b. What type of security architecture or standard do you use (ie; Active Directory)?
- A: MS Active Directory

14. RFP: “Creates filters that include, but are not limited to: enrollment, academic, student, budget, finance, HR, advancement and related functions.” Is this referring to segregating data domains by user role, so only authorized users can see and access data appropriate to them or some other sort of filter?  
A: Yes
15. Does a higher education technical steering group exist today?  
A: Not formally
16. Does a higher education program advisory committee exist today?  
A: Not formally

## **Vendor J**

1. Please list each of the source systems at the three higher education institutions and describe the nature of the data contained in each. If you can provide a data dictionary of elements contained in each of these source systems, please do so.  
A: Please see Q19 on p.3.
2. Page 4 of the RFP, under “Tasks”, states “Using the RI Department of Education’s data warehouse model as a guide, the foundation for the OHE’s warehouse has been created in ER/Studio Data Architect.” Could you share the details of this foundation? Could you articulate the nature of the data, specific elements, entities, and their relationships? Do you have a data dictionary you could share?  
A: Please see Q20 on p.3. The RIDE data dictionary is not applicable to this project.
3. Page 4 of the RFP, under “Tasks”, states “The OHE model incorporates elements from the CEDS postsecondary data model, IPEDS and institution data files.” Could you share the list of both CEDS and IPEDS elements, as well as the institution data files? Could you provide a data dictionary?  
A: Please see the data dictionary available on the RI DataHUB. In addition, you can visit the IPEDS, CEDS and Complete College American sites for more information.
4. Is there a data dictionary, including the name of the data source, the name of the applicable field, the field type, the field length, the valid values, for the data to be loaded?  
A: Please see data dictionary appended to these answers.
5. Is the thought that the current data warehouse be replaced entirely by a new data warehouse, or will the new data warehouse just expand the existing one?  
A: This is a new build
6. What database (RDBMS) does the agency use for its existing higher education data warehouse? Can the proposer assume that that database software can be used for the new data warehouse?  
A: The database developed by Drexel is proprietary. Database software is supplied by RIDE.
7. Is the proposer expected to operationalize the new data warehouse? For example, it is expected to load data into the new data warehouse? If so, how many years of data? What kinds of data?  
A: Yes, that is the expectation for 8 years of data.
8. On page 4 of the RFP, it refers to building a data warehouse in 2006. Which vendor built that data warehouse for RIOHE?  
A: Drexel University.

9. In the same paragraph as above on Page 4 of the RFP, the RFP states, “While this database satisfied a need at the time, new requirements and data requests necessitate a much more robust system. “ Could you specify the new requirements and data requests?  
 A: Any new proposed legislation is accompanied by ad hoc data requests. In addition, initiatives such as Complete College America add to the requirements.
10. In the same paragraph as above on Page 4 of the RFP, the RFP further states that, “Comparable data is required by authorities, including the General Assembly, where funding decisions are made by comparing corresponding figures.” Could you specify the data/figures that need to be compared?  
 A: OHE receives requests weekly. They range from enrollment and transfer information to persistence, remediation and courses taken in STEM-related fields.
11. Also on Page 4 of the RFP, the RFP states that one driver of warehousing the data is “to provide consistent data for annual operating budgets and planning purposes”. If there other drivers, can you elaborate on those?  
 A: TBD
12. Is RIOHE using the same non-SSN, student identifier across all institutions and systems? If not, what are the current person identifiers being used?  
 A: There is a matching process.
13. What product/system does the State use to assign the State Assigned Student Identifier (SASID)?  
 A: This is a RIDE developed system. Each college has a separate identifier which is not a part of the SASID system.

## ETL

14. What ETL tool is being used for the existing higher education data warehouse? Do you have a preference that this ETL tool be used for the new data warehouse?  
 A: Please see Q20 on p.3.
15. Page 5 of the RFP, Deliverable #2, states “Updates data across all systems according to set rules established by a technical advisory group and in line with the annual federal and state data calendars.” Can you expand on and detail the vision here? The nature of requirements for updating source systems from warehouse data, the number and breadth of business rules governing such requirements, the frequency with which these updates are required, and the authentication requirements for doing so, may have significant impact on costs.  
 A: Updates should occur at least monthly. This is a new build so these requirements will need to be developed.

## Reporting

16. On Page 4 of the RFP, under “Tasks”, related to Reporting Requirements, the RFP states “The data warehouse project includes ... design and development of campus-wide data warehouse reporting and extract solutions; and maintenance and support of the data warehouse and reporting development tools environment.” However, on page 5, under “Deliverables”, the Reporting Requirements say “Reporting Requirements: Consult with campus users to determine daily and long-term reporting and analytics requirements. Consult with data warehouse report writers and architects to determine how to model the data and structure the ETLs to streamline report writing.” Could you please clarify what is expected of the vendor related to Reporting requirements?  
 A: The data warehouse should support front-end analytics in the future. For this project, users should be able to extract data using SAS, SPSS, SRSS, CSV and Excel.

17. What reporting/business intelligence tool is being used for the existing higher education data warehouse? Do you have a preference that this tool be used for the new data warehouse?  
A: See answer above.
18. What reports are available to be run against the current data warehouse?  
A: None
19. Is the vendor expected to recreate those reports against the new data warehouse?  
A: N/A
20. If the vendor is expected to develop reports, how many reports is the vendor expected to develop?  
A: N/A
21. Is the vendor expected to operationalize any reports? For example, is the vendor expected to integrate any reports with the agency's security/authentication system?  
A: Not at this time
22. What type of end users would be using the reporting tool?  
A: Internal users with a high degree of technical capacity.
23. In the future vision, is there a desire to align a dashboard tool to the data warehouse?  
A: Yes

## Proposal

24. Page 8 of the RFP, under "Proposal Submission", states "2. A separate sealed Cost Proposal as described above." Page 9, under "Technical Proposal / Required Elements", states "The technical proposal should be 10-20 pages in length, respond to each area of the required elements listed above, and contain a cost proposal using the forms in Appendix A." Please clarify - should the cost proposal be contained within the Technical Proposal, or should it be submitted separately from the Technical Proposal?  
A: They are to be submitted separately at the same time and place.
25. On page 9 of the RFP, under "Proposal Submission, it states "Quality of Key Personnel (including Curriculum vitae) should be included in the Technical Proposal. Does the Curriculum Vitae count towards the 10-20 pages maximum or can that be included in an Appendix?  
A: Appendix
26. What is the preferred format of the proposal? Page 9 the RFP outlines one structure (see "TECHNICAL PROPOSAL/REQUIRED ELEMENTS" below), and on page 10 of the RFP, another structure is outlined (see "Elements of Proposal" below).  
*TECHNICAL PROPOSAL/REQUIRED ELEMENTS*
1. Contractor understanding of the Issues (10 points)
  2. Work Plan (20 points)
  3. Capacity of the Agency Effectively to Administer the Project (20 points)
  4. Quality of Key Personnel (including Curriculum vitae) (20 points)
  5. Cost Proposal (30 points)
- Elements of Proposal:*
1. Overview of your firm's qualifications.
  2. Overview of experience working with post-secondary education data, specifically as it relates to multiple state reporting systems and student information systems.
  3. Realistic work plan for the creation and implementation of a data warehouse platform that syncs data from multiple systems.

4. Resume of key individuals involved in the project and name of point person.
5. Fee estimate, expenses to be incurred, and payment schedule.
6. Three references of clients from the past 2 years for whom you have done work relevant to this assignment.
8. Three relevant work samples.  
A: Use the first format, which incorporates the second.

### **Security and Authentication**

27. What security/authentication system does the agency have in place?  
A: MS Active Directory
28. Is there a single sign-on procedure already in place?  
A: Yes
29. Will the vendor be expected to use that security/authentication system?  
A: Yes
30. Can the vendor use the security/authentication system if it chooses to?  
A: Yes, but the vendor will have to pay any licensing.

### **Implementation and Training**

31. Approximately how many staff members will be trained at each campus?  
A: A limited number to be determined.
32. Do you have a learning management system to house and distribute training materials?  
A: No

### **Other Projects**

33. How is this project related to the State of Rhode Island's Longitudinal Data System project?  
A: It is funded by this grant.

### **Vendor K**

1. General - If DB or ETL software is required, please share how many end users you have today and if you have an expectation of the number of users.  
A: It will be a limited number to be determined.
2. General - As it isn't clear when answers are going to be released to all the questions submitted AST is requesting a 2 week extension of the Due Date. As an answer can change the direction of the proposed solution we will need time to make those changes and as questions are due 2 weeks before the proposal we will need time to accommodate any change in directions.  
A: The due date has been extended.
3. Scope of Work - What platforms, both database and ETL, does RIOHE use for the existing data warehouse?  
A: N/A. See Q20 p.3 for available platforms for the project.
4. Scope of Work - Will the same DB platform be leveraged for this project? If not, has RIOHE identified the new DB platform or are you looking for recommendations?  
A: See Q20 p.3 for available platforms for the project.
5. Scope of Work - Will the same ETL solution be leveraged for this project? If not, has RIOHE identified the new ETL solution or are you looking for recommendations?  
A: See Q20 p.3 for available platforms for the project.
6. Scope of Work - If available, please provide the new requirements that cannot be satisfied by the existing system.

A: TBD

7. Scope of Work - Can you further explain the requested deliverable #4 (Creates filters that include, but are not limited to: enrollment, academic, student, budget, finance, HR, advancement and related functions)?

A: These are filters for experience users who may use SAS or SPSS

8. Scope of Work - In reference to deliverable #6 (Sets up for a connection to an analytics program), please either specify the BI/analytics solution that must be 'connected' or confirm that this deliverable need not be applied to this phase of the project.

A: Only the connection needs to be set up. There is no analytic at this time.

9. Scope of Work - In what ways is the existing data model expecting to expand (i.e. subject areas, dimensions, etc.)?

A: It will expand to include IPEDS and other elements as articulated by the TAG.

10. Scope of Work - Please provide details/specifics around the targeted source systems within scope, including any available software names and versions.

A: Please see Q19, p.3.

11. State's MBE Requirements - Is the State's set aside of 10% a mandatory requirement or a goal for this project?

A: Goal for the project when there are subcontractor opportunities.

12. State's MBE Requirements - Does an MBE partner need to be in place at the time of proposal or upon contract award?

A: The expectation is that the MBE would be an approved RI MBE prior to the submission of the proposal. Please visit [www.mbe.ri.gov](http://www.mbe.ri.gov) for further information.