

Addendum #1

Database and ETL Pipeline Developer

Issue Date: Friday, October 1, 2021

Bids Due: Wednesday, October 13, 2021

The bid opening date for this solicitation has been extended to Wednesday, October 13, 2021 at 1:00PM. Bids must be submitted in accordance with the instructions on page 1 of the original solicitation.

Question	Response
Espire has registered office in US and development centre in India. All development work will be done from our India development Centre. I hope this fine?	The winning vendor would need to be registered to do business in the State of Rhode Island at the time of contract award. That process is run through the State of R.I. Secretary of State's Office. Registering is not required in order to bid.
What is the approximate budget for the project?	The budget for this project is approximately \$100,000.
Are off-shore teams of developers / engineers able to work on this project?	The winning vendor would need to be registered to do business in the State of Rhode Island at the time of contract award. That process is run through the State of R.I. Secretary of State's Office. Registering is not required in order to bid.
Is PPSD open to replacing the existing data warehouse/ODS environment with a solution from a vendor that would include an off-the-shelf data warehouse, new ETL scripts for all needed data sources, along with a reliable process for error identification & reporting?	PPSD previously issued an RFP for these data warehouse services. This RFP is for temporary support to maintain and/or improve existing on premise databases and ETL pipelines.
Is PPSD considering this RFP in conjunction with the earlier District	PPSD previously issued an RFP for these data warehouse services. This RFP is for temporary support

Student Data Warehouse and Dashboard Solution RFP. Or is this an alternative situation? Any clarity the district can provide on the relationship between the two RFPs would be greatly appreciated.	to maintain and/or improve existing on premise databases and ETL pipelines.
Need to Understand about the domain. Means what is the business process is involved with this requirements.	The vendor will need to be flexible and work with multiple stakeholders across PPSD to learn about their data sources and ETL needs.
Need to understand the end goal or Presentation layout.	Most often, the vendor will need to ensure that data from source systems flowing into SFTP folders are subsequently transformed and loaded into SQL Server databases.
Need to understand source and destination data structure. (Is that different destination database or it would be in single database in data warehouse)	There are multiple sources and destinations involved in this project and the vendor will need to review them individually in providing support to PPSD.
Required various sample of Source Data like Flat file, Excel, SQL and (R Language, Python Script Specially). Need to understand Source and Destination Data Transfer Pipeline like (SFTP).	There are multiple sources and destinations involved in this project and the vendor will need to review them individually in providing support to PPSD.
Need to understand the existing ETL tool like SSIS or using other tools.	PPSD does not currently have an ETL tool supporting its ETL pipelines. Rather, a combination of R and Python scripts, as well as SQL jobs, perform most ETL functions.
Need to confirm ETL Tool in which we will be develop pipelines.	PPSD does not currently have an ETL tool supporting its ETL pipelines. Rather, a combination of R and Python scripts, as well as SQL jobs, perform most ETL functions.
Need to confirm who will be create Tableau Dashboard and Other Presentation.	PPSD staff will develop Tableau dashboards leveraging data in its SQL Server databases.
What is the amount of data stored in the current SQL Server-based data warehouse and operational data store (ODS)?	The existing data warehouse contains approximately 165 GB of data. The oldest data sources include information from school year 2007-08.

For Section III, subparagraph C (Parameters), Data Integrations (iii): what APIs, if any, are available for the routine data transfers applications and additional third party vendor applications?	Some vendors have customizable APIs and others have automated data extracts to SFTP.
How many stakeholders are expected to access this SQL Server-based Data?	A limited number of internal PPSD staff require access to the SQL Server databases. A larger number of internal staff will view data through PPSD's Tableau dashboards that draw from these databases.
Is the existing data warehouse centralized? Does it contain data from all of the 41 schools under PPSD?	Yes
What is the total size of the data stored in the data warehouse and what is the expected growth rate over years? (in GBs, TBs).	The existing data warehouse contains approximately 165 GB of data. The oldest data sources include information from school year 2007-08.
When was the data warehouse initially developed? How many years of historical data does the data warehouse has?	The oldest data sources include information from school year 2007-08.
What version of SQL server is currently in place?	SQL Server 2012 11.0.7507.2
How many fact tables and dimension tables are there in total in the current DWH?	The vendor will need to lead a discovery process to assess and diagnose issues with the structure of existing databases. Right now, there are approximately 500 tables across two key databases.
How many users have access to the current data warehouse?	A limited number of internal PPSD staff require access to the SQL Server databases. A larger number of internal staff will view data through PPSD's Tableau dashboards that draw from these databases.
Do you have a documented data governance policy in place?	No; however, PPSD is reestablishing its data governance committee to establish data governance conventions and policies.
Do you currently use or have any ETL / Data integration tools in place?	PPSD does not currently have an ETL tool supporting its ETL pipelines. Rather, a combination of R and Python scripts, as well as SQL jobs, perform most ETL

	functions.
How many technical staff are there in the Office of Innovation and Accountability team?	The Office of Innovation and Accountability has recently been reconstituted and too few staff members are employed to offer a rating. The district is currently hiring for 4-5 new data-analytic roles, including 1 data engineer, 1 role focused on data governance, and at least 1 role with deep expertise in Tableau dashboard development.
What will be the level of participation of the District's technical team in this project?	PPSD staff will work closely with the selected vendor to support the vendor's discovery and execution.
Is the current data warehouse designed based on the Common Education Data Standards (CEDS)? If not, are you looking to align with CEDS to promote interoperability?	No it is not. PPSD previously issued an RFP for these data warehouse services. This RFP is for temporary support to maintain and/or improve existing on premise databases and ETL pipelines.
How is the data quality currently being maintained? Do you have a data quality management tool?	The vendor will need to review and advise on procedures for better managing data quality.
Does PPSD have a documented data quality standards?	The vendor will need to review and advise on procedures for better managing data quality.
Is developing new dashboards as part of this project completely out of the scope?	PPSD staff will develop Tableau dashboards leveraging data in its SQL Server databases.
How many resources do you expect the vendor to deploy in this project?	The vendor should propose the level and structure of support it can provide. The budget for this project is approximately \$100,000.
Is not having educational institution references exclude us from being awarded, considering we have a strong track record in providing Data Warehousing, Data Analytics, Business Intelligence, and MDM services to governmental organizations and private enterprises?	No, as references are only one evaluation criterion of several.
Can the work be performed remotely? Do you expect any work to be	PPSD does not require onsite work.

performed on-site at PPSD's location?	
Is this a fixed bid contract or time & material basis?	The vendor should propose their desired cost structure.
Do you have an incumbent vendor to develop the new data warehouse? Or a new vendor will be selected through a new RFP?	PPSD previously issued an RFP for these data warehouse services. A vendor has not yet been selected. This RFP is for temporary support to maintain and/or improve existing on premise databases and ETL pipelines.
If the new vendor will be selected by releasing a new RFP, will the selected vendor of this project be eligible to participate in that bidding?	PPSD previously issued an RFP for these data warehouse services. A vendor has not yet been selected. This RFP is for temporary support to maintain and/or improve existing on premise databases and ETL pipelines.
Do you have any plan to shift the DWH from on-prem to cloud in the near future?	Yes, but this RFP is for temporary support to maintain and/or improve existing on premise databases and ETL pipelines.
Is there any incumbent serving in this project?	No.
How many ETL Developers were involved in maintenance?	Several PPSD staff members developed R and Python scripts, as well as SQL jobs, to facilitate ETL over the past several years.
Is the existing Data Warehouse custom built?	Yes.
Who implemented the existing Data Warehouse solution?	Several PPSD staff members developed the SQL Server databases over the past several years.
Who is the incumbent maintaining the Data Warehouse and ETL?	No.
Is the ETL developers need to work on premise?	PPSD does not require onsite work.
Is remote work allowed?	PPSD does not require onsite work.
Is there any mandatory requirements for the ETL developers to be present in person?	PPSD does not require onsite work.

Is vendor for building new data warehouse identified?	No, this vendor has not yet been selected.
Can we recommend & use open source ETL tools?	Yes.
What is the estimated budget for this project?	The budget for this project is approximately \$100,000.
What is the physical nature of the Data Sources? Are they all SQL Server?	Data sources exist within third party vendor applications (accessible via API and SFTP exports), in shared network drives as flat files, and in SQL Server databases.
Are you open to using a different ETL Product. In this case, ours - Relational Junction.	Yes.
Other than via Tableau Server, how is data in the data warehouse and ODS expected to be accessed or used?	PPSD staff will also directly query the databases as needed.
How many users need to access the data in the data warehouse and ODS via Tableau? How many via other methods?	Tableau dashboards will accessed by approximately 100 central office and school leadership staff members. Access to the underlying databases will only be required for a small number of district data staff (5-7).
What is the volume of data stored in your current data warehouse and ODS?	The existing data warehouse contains approximately 165 GB of data. The oldest data sources include information from school year 2007-08.
How many data tables exist in your current data warehouse and ODS?	The vendor will need to lead a discovery process to assess and diagnose issues with the structure of existing databases. Right now, there are approximately 500 tables across two key databases.
Do you have a preferred project staffing? That is, do you have any expectations around the number of FTEs? If so, what are your expectations?	The vendor should propose the level and structure of support it can provide. The budget for this project is approximately \$100,000.
Do you expect any internal resources to support or engage in more than an advisory or consulting role with the project (other than your internal project lead or primary point of contact)? If so, how do you expect them to support or	PPSD staff will work closely with the selected vendor to support the vendor's discovery and execution.

engage with the project?	
In similar scenarios as described in your RFP where we are unsure of the durability of existing systems and processes or there is significant variability in systems and processes, we often recommend implementing a new cloud-based data warehouse in parallel to the existing on-premise solution. This parallel system can provide security, standardization, reliability, availability, and predictability. As we correct or improve data pipelines and integrations, they would not need to be rebuilt later. They would have a greater degree of reliability without interfering with anything that is working in the existing infrastructure. We would of course consider the suitability of alternatives during our discovery and consult with your stakeholders prior to making a final recommendation. If we submit a proposal, we will likely include this in our approach to reduce project costs and enable the most efficient and predictable path towards the necessary improvements in your data infrastructure. Would you be open to considering such an approach?	Yes, however PPSD previously issued an RFP for these data warehouse services. This RFP is for temporary support to maintain and/or improve existing on premise databases and ETL pipelines.
Section IV describes a collaboration with PPSD to support the development of a new data warehouse through an engagement with another vendor starting in April 2022. Relative to that vendor: - What role would that other vendor have (e.g. platform/software vendor such as Microsoft SQL Server, Snowflake, AWS Redshift; or a service provider; or another type of vendor)? - Is any of that determined now? Or is a vendor selected? If so, what vendor or vendor system is selected?	Prospective vendors for the other RFP include a combination of technical services providers, as well as providers with product offerings. A vendor has not yet been selected.

What roles/types of resources does PPSD expect to be part of the vendor's proposed project team?	PPSD staff will work closely with the selected vendor to support the vendor's discovery and execution.
How many candidates can we propose for the Database and ETL Pipeline Developer position?	The vendor should propose the level and structure of support it can provide. The budget for this project is approximately \$100,000.
Can the services for this project be provided remotely? If yes, does the resource need to reside in the US?	PPSD does not require onsite work. The winning vendor would need to be registered to do business in the State of Rhode Island at the time of contract award. That process is run through the State of R.I. Secretary of State's Office. Registering is not required in order to bid.
What is the budget for this project?	The budget for this project is approximately \$100,000.
Is there an incumbent vendor currently providing similar services?	No.
Is there a standard Pricing Proposal format that PPSD could share for reference?	No, please use whatever format you feel most clearly articulates your proposed pricing structure.