

<b>STATE OF NORTH CAROLINA</b>	<b>REQUEST FOR INFORMATION NO.88-260027-JE</b>
Central Piedmont Community College	Due Date: May 12, 2026 by 2pm
<i>Refer <b><u>ALL</u></b> Inquiries to:</i> Jennifer Ennis	Issue Date: April 14, 2026 Commodity: 811122 – Data Lake Platform – Software as a Service
E-Mail: Jennifer.ennis@cpcc.edu	Using Agency Name: Central Piedmont Community College

**SUBMITTAL INSTRUCTIONS:** Responses to this Request for Information (RFI) shall be submitted via the electronic vendor portal (eVP) by the due date and time specified herein. Additional information can be found at <https://evp.nc.gov/>.

**NOTICE TO VENDOR**

Request for Information (RFI) responses will be received electronically via the electronic Vendor Portal (eVP) until May 12, 2026 by 2pm on the day of opening and then opened. Additional information can be found at the eVP updates for Vendors link: <https://eprocurement.nc.gov/news-events/evp-updates-vendors>

**EXECUTION**

VENDOR NAME:	E-MAIL:	
STREET ADDRESS:	P.O. BOX:	ZIP:
CITY & STATE:	TELEPHONE NUMBER:	TOLL FREE TEL. NO:
TYPE OR PRINT NAME & TITLE OF PERSON SIGNING:	FAX NUMBER:	
AUTHORIZED SIGNATURE:	DATE:	

## 1.0 EXECUTIVE SUMMARY

Central Piedmont Community College (Central Piedmont) is actively evolving its data and analytics capabilities to better support institutional reporting, operational decision making, advanced analytics, and future data driven initiatives. As part of this evolution, Central Piedmont is seeking to better understand available Data Lake platforms, architectures, and capabilities that can operate effectively within a hybrid environment, integrate with existing enterprise systems, and align with established governance, security, and operational standards.

This Request for Information (RFI) is intended to collect information and recommendations regarding: modern Data Lake solutions that can support the College's current and future data, analytics, and integration needs. Including but not limited to:

- Architecture and platform capabilities, particularly in hybrid and cloud based environments.
- Data ingestion, integration, and management approaches for diverse data sources.
- Governance, security, and compliance features suitable for enterprise and public sector environments.
- Operational considerations, including monitoring, support, and long term sustainability.
- Scalability, performance, and cost management characteristics.
- Vendor experience and roadmap relevant to organizations of similar size and complexity.

Central Piedmont requests detailed point-by-point responses showing how your firm would address the items in Section four (4) of this RFI:

## 2.0 RFI PROCEDURES

### 2.1. SCHEDULE

Respondents will have four (4) weeks to prepare their submissions to this RFI. Responses must be received by the date, time, and location specified on the cover sheet of this RFI.

There is no guarantee that all RFI submissions will be invited to present. If selected, respondents will be notified of the specific date and time approximately two weeks in advance of their presentation date.

#### Intended Schedule of Events

All times listed are EST.

April 14, 2026	RFI issued
April 24, 2026 by 2pm	Written questions due from vendors
April 29, 2026	Response to Written Questions (via Electronic Vendor Portal (eVP))
May 12, 2026, by 2pm	Submission of Request for Information responses due

### 2.2 RFI RELATED QUESTIONS/CLARIFICATIONS

Any questions regarding this RFI should be emailed on or before April 24, 2026 at 2:00 pm EST to the attention of: jennifer.ennis@cpcc.edu. All questions must be submitted in writing. Please enter "Questions RFI 88-260027-JE" as the subject for the email. An addendum containing any general clarification questions and their answers will be issued as an addendum to this RFI within the electronic vendor portal (eVP).

Vendors responding to the RFI shall designate a single contact within that company for receipt of all subsequent information regarding this RFI.

### 3.0 RFI RESPONSE

Central Piedmont recognizes that considerable effort will be required in preparing a response to this RFI. **However, please note this is a request for information only, and not a request for services.** The Vendor shall bear all costs for preparing this RFI. Responses for this RFI will be used to:

- Understand current market capabilities and architectural approaches for enterprise Data Lake solutions.
- Assess how different solutions align with Central Piedmont’s technical, governance, and operational context.
- Identify strengths, limitations, assumptions, and tradeoffs associated with various approaches.
- Inform future planning, architecture decisions, and potential procurement activities.

#### 3.1 **CONTENT AND FORMAT**

Central Piedmont expects concise, detailed, point-by-point responses to each of the RFI response items identified in Section 4.0 of this RFI. Central Piedmont is not interested in brochures or “boilerplate” responses. Instead, responses should clearly define how the vendor’s proposed solution(s) would meet Central Piedmont’s environment, operating model, and governance approach. Any issues or exceptions to Central Piedmont’s defined environment, operating model, and governance should also be identified and explained.

Where appropriate, the response should also include annotated network drawings showing where each of the pieces of equipment in the proposed solution would be located and how those devices would be interconnected.

The response should define all services that would be required by the proposed solution. The response should also include:

- The vendor’s understanding of the project and services by addressing Central Piedmont Community College’s environment, operating model, and governance approach;
- An estimated total cost of ownership for the solution including continued compliance with emerging industry standards.

#### 3.2 **MULTIPLE RESPONSES**

Multiple responses can be provided from a single vendor provided that each response is comprehensive, addresses all Central Piedmont’s defined environment, operating model, and governance, and is truly unique. Please clearly mark responses as “Response #1, Response #2, etc.

### 4.0 REQUESTED INFORMATION FROM VENDORS

Central Piedmont Community College is interested in hearing from qualified vendors to describe their solutions clearly and concisely, focusing on how their offerings would fit within Central Piedmont’s environment, rather than providing generic marketing material.

Responses should be tailored to Central Piedmont's defined environment, operating model, and governance approach as described in this RFI.

Central Piedmont Community College currently operates within the following:

- Hybrid environment with on-prem systems and Azure workloads, governed via a centralized Azure Landing Zone architecture.
- A hybrid (hub and spoke) Data & Analytics operating model, with centralized governance and distributed analytics capabilities.
- Multiple data domains (administrative, academic, operational, security, etc.) with varying sensitivity and compliance needs.
- Strategic goals: analytics, reporting, self-service BI, advanced analytics/AI, and long-term data retention.

## **Assumptions and Clarifications**

- Central Piedmont operates within a hybrid environment leveraging Azure Landing Zones; vendors should assume alignment to standard Azure governance practices
- The College uses a centralized governance model with distributed analytics capabilities across business units
- Institutional data is classified and governed based on sensitivity, and solutions should support classification-driven controls
- Data governance roles (e.g., Data Owners, Data Stewards, Data Custodians) are expected to participate in approval, certification, and oversight processes
- Responses should assume integration with existing enterprise identity, security monitoring, and data platforms
- Where specific details are not provided, vendors should state assumptions made in their response

Respondents are invited to provide information on the following subjects:

### **4.1 ARCHITECTURE & PLATFORM CAPABILITIES**

- a. Describe your reference architecture for a Data Lake in a hybrid Azure environment. How does it integrate with existing Azure Landing Zones?
- b. How does your solution support separation of environments (dev/test/prod) and multi-subscription or multi-account models?
- c. What storage formats are supported (e.g., open formats vs. proprietary)? How is schema evolution handled?
- d. How does the platform support structured, semi structured, and unstructured data at scale?
- e. Can the solution evolve from a traditional data lake into a lakehouse or analytics optimized architecture over time?
- f. Describe your implementation methodology for enterprise data platforms, including discovery, design, build, test, and transition phases.

### **4.2 DATA INGESTION & INTEGRATION**

- a. What ingestion methods are supported (batch, streaming, CDC, API based)?
- b. How does the solution integrate with on-premises systems and legacy platforms?
- c. Describe how data validation, error handling, and retries are managed during ingestion.
- d. Can ingestion pipelines be reused and templated across domains?
- e. How does the solution align with standardized integration and SOP practices already in place?
- f. Describe how the platform supports data classification and sensitivity tagging. Include how classification labels can be applied, inherited, or enforced and how they influence access control, masking, or other protection mechanisms.

- g. Describe how the platform supports an enterprise data catalog and a data glossary. Include capabilities for maintaining standardized definitions, enabling dataset discovery, identifying trusted or certified datasets, and associating business metadata with technical metadata.
- h. Describe how the platform validates data during ingestion and transformation processes. Include capabilities for schema validation, required field enforcement, format validation, and detection of incomplete or inconsistent records. Describe how validation failures are handled, logged, and notified to administrators or data stewards.

#### **4.3 GOVERNANCE, SECURITY & COMPLIANCE**

- a. How does the platform enforce data governance (metadata, lineage, classification)?
- b. Describe role-based access control and integration with enterprise identity providers.
- c. How is data encryption handled at rest and in transit?
- d. How does the solution support policy driven controls aligned with Azure Policy and centralized governance models?
- e. What compliance frameworks or benchmarks does the platform support (without requiring customization)?
- f. Describe how the platform supports ongoing data quality improvement and governance. Include capabilities for defining quality rules, monitoring quality metrics, providing stewardship visibility into data quality issues, and managing remediation workflows when quality thresholds are not met.
- g. Describe how the platform supports governance and transparency for advanced analytics and AI use cases. Include capabilities for dataset traceability, model governance, explainability, and controls that help ensure sensitive or restricted data is used appropriately in analytical or AI-driven workloads.
- h. Describe how the platform supports institutional data governance roles, including Data Owners, Data Stewards, and Data Custodians. Include capabilities that allow owners/stewards to review, approve, certify, or restrict datasets and how stewardship accountability is maintained across data domains.
- i. Describe how the platform supports compliance with regulatory and legal obligations relevant to higher education. Include capabilities that support data protection, audit reporting, retention management, and evidence generation for regulatory reviews

#### **4.4 DATA MANAGEMENT & LIFECYCLE**

- a. How is data lifecycle management handled (hot/warm/cold/archive tiers)?
- b. Can retention policies vary by data domain or classification?
- c. How does the platform support data quality rules and profiling?
- d. Describe mechanisms for data versioning and historical snapshots.

#### **4.5 ANALYTICS, BI & ADVANCED USE CASES**

- a. How does the solution integrate with BI and analytics tools (self-service and enterprise)?
- b. What support exists for advanced analytics, ML, or AI workloads using lake data?
- c. Can different personas (analysts, data scientists, business users) safely access the same data?
- d. How does the solution prevent “data duplication” across analytics use cases?

#### **4.6 OPERATIONS, MONITORING & SUPPORT**

- a. How is monitoring and alerting handled across ingestion, storage, and access?
- b. What operational metrics are exposed to administrators?
- c. Describe backup, recovery, and disaster resilience options.
- d. What is the vendor’s support model, escalation path, and SLAs?
- e. Describe how the platform supports monitoring and detection of suspicious or unauthorized data access. Include capabilities for audit logging, anomaly detection, integration with security monitoring platforms (e.g., SIEM), and alerting mechanisms for potential data misuse or exfiltration.

- f. What common failure modes have you observed in higher-education data lake initiatives, and how does your approach prevent them.

#### **4.7 SCALABILITY, PERFORMANCE & COST MODEL**

- a. How does the platform scale with increasing data volume and concurrency?
- b. What performance tuning options exist?
- c. Describe the pricing model and key cost drivers.
- d. How does the solution help prevent uncontrolled cost growth?

#### **4.8 VENDOR EXPERIENCE & ROADMAP**

- a. Describe your experience supporting public sector or higher education organizations.
- b. Provide examples of organizations with similar scale and governance needs.
- c. What is your product roadmap for data lake and analytics capabilities?
- d. How do customers influence roadmap priorities?

#### **4.9 EXIT STRATEGY & PORTABILITY**

- a. How can data be exported or migrated if the organization changes platforms?
- b. Are data formats and metadata portable?
- c. What dependencies would remain after exit?

#### **4.10 KNOWLEDGE TRANSFER**

- a. Describe your approach to knowledge transfer to internal staff during and after implementation.
- b. What artifacts (documentation, patterns, reusable pipelines, governance assets) are delivered to ensure long-term internal sustainability.