Duty Statement

☐ Current   ☒ Proposed

Classification Title	Division
Information Technology Specialist III	Data Infrastructure
Working Title	**Office/Unit/Section**
Senior Data Engineer	Office of Cradle-to-Career Data
Position Number	**Effective Date**
424-100-1415-001	
Name	**Date Prepared**
	2023-08-08

General Statement

The Cradle-to-Career (C2C) program is a longitudinal data system that will provide policymakers, researchers, educators, students, families, and other stakeholders' answers to key questions about student progression and outcomes. Ultimately, C2C will be a useful tool that will help provide critical information about the pipelines from early care to K-12 to higher education skills training and employment along with health and human services data. This data system will help support teachers, advisors, parents, and students and be an evidence-based tool that decision makers and researchers can use to help California adopt more equitable policies by providing insight into how educational experiences impact students' subsequent academic achievement, work, and earnings.

Under the administrative direction from the Director of Data Infrastructure, the Information Technology Specialist III (ITS III) serves as the Senior Data Engineer (SDE) and lead on the development of data pipelines and developer on the modern data stack for the Cradle-to-Career Data System. The Senior Data Engineer (SDE) has an advanced level of understanding and technical expertise to develop standards, analytics and methods in which data is stored, arranged and integrated. The incumbent will ensure data is available as a service for data

warehouses, digital services, and dashboards. The Senior Data Engineer will plan, develop, and document policies and best practices for the Office of Cradle-to-Career Data. The SDE will provide direction and assistance to systems analysts, IT cloud engineers, and other systems architects. The incumbent will support the design and development of the cloud data architecture, monitor and report system health and resource consumption to the C2C leadership team. The SDE will work closely with external vendors, internal business groups, and C2C management regarding the implementation and operation of the C2C Data System.

Job Functions[1]

Design and Develop   45% (E)

- Leads data pipeline engineering to develop the infrastructure and processes necessary to support data as a service, including leading the design of systems infrastructure and architecture, documenting configuration, change management aspects of the defined solutions, reviewing performance and security reports and modifying systems to address performance, security, capacity planning and maintenance issues.

- Supports the design, implementation, maintenance and operations of the C2C Data System using technologies including, but not limited to, the following: Amazon Web Services (AWS), cloud networking, Snowflake, Tamr, and Ping Identity.

- Plans, designs, develops, and implements processes for C2C Master Data Management, including, but not limited to, the following: data intake, validation, quality, modeling, and visualization.

- Ensures data moves securely and performantly to enable data for consumption by data services.

- Analyzes pipeline throughput, issues, and other metrics to inform and implement improvements to the automation platform.

- Standardizes and improves existing pipelines to adhere to best industry standard data practices.

- Oversees data engineering work of external consultants to ensure alignment with business needs and stakeholders.

- Advise management on formulating data engineering strategy, policy, and governance.

- Collaborates with business and management teams to establish business data needs and develops strategies for long-term data platform architecture.

---

[1] Functions: Essential (E) or Marginal (M)

- Develop, plan, review, document, advise, and communicate data engineering policies, processes, procedures, standards, guidelines, and best practices to project team members.

Data Quality and Security   20% (E)

- Design, develops, and documents patterns and practices to help ensure all data is profiled, reviewed, and accurate to support critical decisions.
- Designs and develops data quality metrics and data quality alerting system to ensure integrity and validation of the data for the Cradle-to-Career Data System.
- Act as a subject matter expert to develop high-quality, well-tested, maintainable code across an entire data engineering lifecycle, using modern software development best practices, including continuous integration and continuous delivery.
- Design, develops, and maintains documentation so all Office of Cradle-to-Career Data team members know their roles and responsibilities in meeting data quality metrics.

Collaborate and Support   20% (E)

- Collaborates with Data Infrastructure members and contractors, supports and implements data engineering platforms for the Cradle-to-Career Data System to meet the business needs.
- Collaborates with Data Infrastructure team members, to guide and implement architectural improvements to the data infrastructure.
- Collaborates with Data Infrastructure team members to evolve Continuous Integration and Continuous Delivery strategy, including instrumenting robust monitoring, altering and recovery.
- Collaborates with external data providers, both internal and external to the State to define requirements for data pipelines to securely ingest data into the Data System.
- Supports Office of Cradle-to-Career Data team members to identify and implement modeling approaches and methods.

Share and Consult   10% (E)

- Conducts presentations to share documented patterns and best practices utilized by the Cradle-to-Career Data System through a variety of forums including communities of practice.
- Advise on data engineering requirements and needs as part of project solicitation, vetting, and scoping processes.
- Maintains and improves knowledge of emerging technologies for various cloud environments, Identity Management solutions, Master Data Management technologies and Visualization tools, state accessibility requirements, develops staff skill sets and maintains the stability of the Data System.

Marginal Functions   5% (M)

- Performs a variety of administrative activities that provide efficiencies to C2C and the Data Infrastructure team. Duties include, but are not limited to, the following: workload management and participation in special ad hoc committees, staff meetings, teams and projects, time reporting and procurements. Prepares correspondence, reports, and issue papers related to data and software, hardware, and C2C system.

Supervision Received

The Senior Data Engineer will report to the Director of Data Infrastructure.

Supervision Exercised

None, however, provides technical and project management leadership to teams within the department as well as external stakeholders, and has defined responsibility and authority for decision making related to projects, or in an advisory function.

Personal Contacts

The incumbent will work with teams across the Office of Cradle-to-Career Data as well as GovOps Agency staff, community organizations, a wide range of stakeholders, and external contractors and advisors.

Actions and Consequences

The incumbent's duties are imperative to the successful implementation of the Data Infrastructure for the C2C program. Failure to deliver this system will be failure to the C2C program mission to foster evidence-based decision-making to help Californians build more equitable futures and empower individuals to reach their full potential.

Functional Requirements

The demands described here are representative of those that you must meet, with or without a reasonable accommodation, to successfully perform the essential functions of the job. The job requires flexibility in work hours and work overtime occasionally, including night and/or weekend hours, as necessary. Requires daily use of a personal computer and related software applications at a workstation. Requires ability to complete tasks that typically may require making repetitive hand movements in the performance of daily duties, with or without reasonable accommodations and modifications to facilitate such tasks. Requires occasional travel to attend meetings and hearings.

Working Conditions

The employee regularly works in an indoor and climate-controlled office setting under artificial light. The employee's workstation is located in Sacramento, CA, and is equipped with standard or ergonomic office equipment, as appropriate. May sit for an extended period using a keyboard and video display terminal. Based on departmental or operational needs, work can be performed remotely. The employee can work from anywhere within California. Occasional travel may be required to attend offsite meetings, conferences, and training classes. On occasion, may require flexible work schedules, including some evening hours to complete assignments, meet deadlines, and provide support to the C2C leadership team.

Attendance

Must maintain regular and acceptable attendance at such a level as is determined at C2C's sole discretion. Must be regularly available and willing to work the hours the department determines are necessary or desirable to meet its business needs.

Background Checks and Clearance

The successful candidate will be required to pass a criminal background check (see Education Code 10873).

Desirable Qualifications

- Bachelor's degree in information technology, social science, data analysis, economics, public policy, education policy, or a related field.
- Current AWS Data Analytics certificate.
- Current AWS Solution Architect Associate certificate.
- Experience designing, developing, deploying, securing, and maintaining infrastructure of cloud-based data systems.
- Experience integrating disparate data systems, preferably those housing educational data.

- Experience designing, implementing, and optimizing cloud-based data warehouse (e.g., Snowflake).

- Experience designing and implementing Master Data Management (MDM) solutions (e.g., Tamr).

- Experience designing and implementing industry standard Identity and Access Management (e.g., Ping Identity).

- Demonstrated experience designing and developing accessible, accurate data visualizations from large datasets.

- Experience monitoring systems stability and performance and ensure system availability, reliability, and usability.

- Experience managing agile teams of professionals and effectively coaching them to achieve high performance.

- Experience at the senior leadership level of planning and managing strategic projects and being accountable for results.

I have read and understand the duties listed above and I can perform these duties with or without reasonable accommodation. (If you believe reasonable accommodation is necessary, discuss your concerns with the hiring supervisor. If unsure of a need for reasonable accommodation, inform the hiring supervisor, who will discuss your concerns with the Personnel analyst.)

Duties of this position are subject to change and may be revised as needed or required.

Employee Name: _____

Employee Signature: _____

Date: _____

I have discussed the duties of this position with and have provided a copy of this duty statement to the employee named above.

Supervisor Name: _____

Supervisor Signature: _____

Date: _____