Data Platform Lead (DBRE) for English Speaker

English speaking role

Job Information

Recruiter
PROGRE Ltd

Hiring Company
AI company

Job ID
1619403

Industry
Internet, Web Services

Company Type
Small/Medium Company (300 employees or less)

Non-Japanese Ratio
About half Japanese

Job Type
Permanent Full-time

Location
Tokyo - 23 Wards, Minato-ku

Salary
7 million yen ~ 11 million yen

Work Hours
9:00 ~ 18:00

Holidays
土日祝日、年末年始および会社指定日

Refreshed
November 10th, 2026 02:00

General Requirements

Minimum Experience Level
Over 3 years

Career Level
Mid Career

Minimum English Level
Business Level (Amount Used: English usage about 75%)

Minimum Japanese Level
Daily Conversation Level

Minimum Education Level
Bachelor's Degree

Visa Status
Permission to work in Japan required

Job Description

Position Overview

Position Overview

As the product and its data models have expanded alongside rapid business growth, the team's mission is to establish sustainable, continuously compounding data governance capabilities—not simply conduct a one-time review.

As Data Platform Lead, you will define the team's technical direction and roadmap and drive execution. This role goes beyond database operations and tuning: you will lead data-model redesign and establish a decision framework for allocating data between TiDB and Snowflake.

Key Challenges

- Define data lineage and tiering policies across a large codebase and entities spanning Aurora MySQL, TiDB, Aurora PostgreSQL (for some services), and Snowflake.
- Standardize performance and design decisions using measurements and lineage, supported by benchmarks, reliable tests, data classification, and architecture decision records (ADRs).
- Redraw the boundary between OLTP and analytics by moving heavy aggregation workloads to Snowflake.
- Lead a multi-quarter redesign of core customer and segment data models.
- Establish proof that changes preserve system behavior before beginning major refactoring or migrations.

Responsibilities

Hands-on DBRE Responsibilities

- Improve Aurora MySQL and TiDB performance and capacity through query-plan optimization, lock mitigation, and large-table schema changes using tools and approaches such as gh-ost, pt-online-schema-change, and Online DDL, guided by benchmarks.
- Design and prioritize the governance foundation, defining minimum viable standards for lineage, test reliability, benchmarking, data classification, ADRs, local reproduction environments, and collaboration with external experts.
- Design and lead workload-reduction initiatives, including retention-based deletion, event attribute modeling, offloading heavy aggregations to Snowflake, and reverse ETL.
- Lead incident response from coordination through root-cause analysis (RCA) and implementation of preventive measures.
- Personally design and implement solutions in the most challenging areas and establish engineering standards through code.

Technical Leadership Responsibilities

- Maintain and update the 12-month technical roadmap, reviewing it monthly and refreshing phase definitions, exit criteria, dependencies, and critical paths quarterly.
- Establish a decision framework for determining which entities remain in TiDB and which move to Snowflake, based on data classification and lineage.
- Lead the next-generation redesign of customer and segment data models, following a profiling-first approach and maintaining an evidence trail from classification, lineage, and benchmarks through decisions and documentation.
- Build consensus with business stakeholders on changes such as data deletion under retention policies and migration of segment definitions; explain impacts and risks in non-technical language.
- Establish and operate collaboration processes with database and data-warehouse vendors, including regular reviews, a single point of contact, escalation paths, and documented decisions.
- Coach and onboard team members and participate in hiring. The scope of people management can be discussed based on your preferences and the organization's needs.

What This Role Is—and Is Not

This is not a role limited to rewriting shared libraries or short-term database tuning. Nor is it complete once data has been moved to a data warehouse. The core responsibility is to build repeatable governance and verification capabilities that make large-scale model changes safe and sustainable.

Team & Ways of Working

- You will join the Data Platform team within Mico Engage AI and report to the Senior Tech Lead.
- The current team includes three senior members in Japan and three members in India.
- The team works across Japanese and English. Daily stand-ups are shared asynchronously in Japanese and English via Notion.
- The team works with database and data-warehouse vendor experts through regular reviews.
- Development teams use Claude (Anthropic) as part of their development workflow.

Technology Stack

- Application: TypeScript / Node.js (monorepo)
- Databases: Aurora MySQL, TiDB, and Aurora PostgreSQL for some services
- Data warehouse: Snowflake
- Cloud: AWS

Required Skills

Required Qualifications

- Approximately 3+ years of hands-on ownership of production MySQL-family RDBMS design and operations (e.g.,

Aurora MySQL, TiDB, or MySQL), plus approximately 2+ years making design decisions as a Tech Lead or Architect.
- Strong operational expertise in MySQL-family databases, including query-plan optimization, lock mitigation, and large-table DDL operations (e.g., gh-ost, pt-osc, or Online DDL), with the ability to quantify capacity and scaling limits.
- Practical relational data-modeling experience, including entity design, normalization/denormalization, index design, and transaction-boundary design.
- Experience designing and leading a data-model or schema migration in a large live system without service downtime, including migration plans involving dual writes, backfills, and gradual cutovers.
- Experience setting engineering/design standards and embedding them in a team's decision-making practices.
- Fluent or higher proficiency in either English or Japanese across speaking, writing, and reading, plus at least conversational proficiency in the other language.

Preferred Qualifications

- Experience defining the boundary between OLTP and analytics systems, including offloading aggregation workloads to a DWH such as Snowflake and implementing reverse ETL.
- Production-level TypeScript / Node.js development skills and experience working in or refactoring a large monorepo or shared libraries.
- Experience with SLOs, incident leadership, data classification, access controls, and retention policies.
- Experience leading technical selection or investment decisions across stakeholders, quantifying trade-offs, and gaining agreement from non-engineering partners.
- Experience mentoring engineers and contributing to hiring processes, interview design, or candidate experience.
- Basic knowledge of PostgreSQL.
- Ability to lead technical discussions and write technical documentation in English.
- Deep knowledge of distributed SQL databases such as TiDB, including capacity planning, garbage collection (GC), region design, and performance limits.
- Experience leading Snowflake or another DWH implementation, including cost management and defining system responsibilities.
- Experience implementing data governance practices such as data catalogs/classification, lineage, ADRs, data contracts, and privacy or audit requirements under applicable laws such as Japan's APPI or GDPR.
- Experience defining and operationalizing data-retention and deletion policies with business stakeholders.
- Experience rebuilding test strategies, improving test reliability, or leading large-scale system modernization.
- Experience designing vendor collaboration models and reducing infrastructure costs with clear business-level explanations.
- Contributions to open-source projects, technical publications, or conference presentations.

Company Description