CGI

Conversational Database Interface

Transform Data Access Through Natural Language

Executive Summary

In today's data-driven business environment, organizations face a critical challenge: valuable insights remain locked within databases, accessible only to technical teams with SQL expertise. This creates bottlenecks, slows decision-making, and limits the democratization of data across the enterprise.

CGI (Conversational Database Interface) eliminates this barrier by enabling non-technical users to interact with corporate databases using natural language. Ask questions in plain English, receive instant answers with visualized data—no SQL knowledge required.

The Business Challenge

Data Accessibility Gap

Business users wait hours or days for technical teams to extract data, creating decision-making delays and operational inefficiencies.

Rising Analytical Costs

Organizations spend significant resources on data analysts performing routine queries that business users could execute themselves.

Security Concerns

Granting direct database access to non-technical users risks accidental data corruption, deletions, or security breaches.

Knowledge Silos

Critical business intelligence remains trapped within databases, inaccessible to stakeholders who need it most.

The CGI Solution

CGI leverages advanced artificial intelligence to bridge the gap between business users and enterprise data. Our conversational interface interprets natural language questions, generates secure SQL queries, and delivers actionable insights in seconds—not hours.

90%

Reduction in Query Wait Time

100%

Read-Only Security

0

SQL Training Required

Four Strategic Imperatives for Data Democratization

CGI delivers value across four key imperatives, transforming how your organization interacts with data while maintaining enterprise-grade security and governance.

Envision

Natural language data access strategy aligned with business objectives

Explore

Evaluate ROI-led use cases across departments and functions

Engineer

Build secure, scalable foundations with read-only architecture

Expand

Scale to enterprise-wide adoption with governance and compliance

Core Capabilities

Security-First Architecture

CGI implements defense-in-depth security with multiple validation layers. All database interactions are strictly read-only—no modifications, deletions, or schema changes are permitted. Multi-layered SQL injection protection, prompt injection defense, and comprehensive audit logging ensure enterprise-grade security.

- **Read-Only SQL Enforcement:** Only SELECT queries execute; all write operations blocked at multiple layers
- **SQL Injection Protection:** 22+ test scenarios validated; keyword filtering and query normalization prevent bypass attempts
- **LLM Prompt Injection Defense:** Validation after AI generation ensures security boundary integrity
- **Audit Trail:** Complete query logging for compliance and forensic analysis
- **Credential Management:** System keyring integration and Docker Secrets support

Natural Language to SQL

Users ask questions in plain English. CGI's AI engine interprets intent, generates optimized SQL queries, and presents results in human-readable formats. No technical knowledge required.

Example: "Show me top 5 customers by spending this quarter" → Instant results with visualization

📚 Knowledge Base Integration (RAG)

Beyond database queries, CGI leverages Retrieval-Augmented Generation (RAG) with vector similarity search to answer questions about company policies, procedures, and documentation stored in your knowledge base.

Example: "What is our return policy?" → Retrieves and synthesizes policy documentation

Conversational Memory

CGI maintains session context, understanding follow-up questions and referencing previous queries. This enables natural, flowing conversations rather than isolated question-answer pairs.

Business Value Proposition

Quantifiable Benefits

Operational Efficiency

Reduce data analyst workload by 60-80% for routine queries. Business users self-serve, freeing analysts for high-value work.

Faster Decision Making

Access insights in seconds instead of hours or days. Accelerate strategic and tactical decisions with real-time data access.

Cost Reduction

Minimize headcount requirements for data extraction teams. Reduce training costs—no SQL courses needed.

Risk Mitigation

Eliminate accidental data corruption. Read-only architecture prevents costly database incidents.

Industry-Agnostic Applications

Sales & Marketing

- "Show me conversion rates by marketing channel this month"
- "Which products have the highest customer satisfaction scores?"
- "List customers who haven't purchased in 90 days"

Finance & Operations

- "What are our top 10 expenses by category this quarter?"
- "Show inventory levels for products below reorder threshold"
- "Calculate average order processing time by region"

Customer Support

- "How many support tickets are open for product X?"
- "What is our average response time by priority level?"
- "Show customer satisfaction trends over the last 6 months"

Executive Leadership

- "Display revenue by business unit year-over-year"
- "What are our key performance indicators this quarter?"
- "Show employee productivity metrics by department"

Competitive Differentiation

CGI stands apart from traditional business intelligence tools through:

- **Zero Configuration BI:** No complex dashboards or pre-built reports required
- **Ad-Hoc Analysis:** Answer questions you didn't know you needed to ask
- **Security by Design:** Read-only architecture prevents data corruption from day one
- **Rapid Deployment:** Docker-based infrastructure deploys in minutes, not months
- **Model Agnostic:** Works with any OpenAI-compatible LLM—on-premises or cloud
- **Open Architecture:** Integrates with existing PostgreSQL databases without migration

Solution Architecture

CGI leverages proven enterprise technologies in a containerized, cloud-ready architecture designed for scalability, security, and reliability.

Technology Stack

PostgreSQL + pgvector

Industry-standard relational database with vector similarity search for RAG capabilities. Scales from small teams to enterprise deployments.

Large Language Models

OpenAI-compatible API supports cloud providers (OpenAI, Anthropic) or on-premises deployment (llama.cpp, Ollama) for air-gapped environments.

Python Application Layer

Modular architecture with separate concerns: database management, LLM client, and interactive CLI interface.

Docker Containerization

Fully containerized deployment ensures consistency across development, staging, and production environments.

Implementation Approach

Phase 1: Pilot Deployment (Weeks 1-2)

- Deploy CGI on development/test database with sample data
- Onboard 5-10 pilot users from key business units
- Gather feedback on query accuracy and user experience
- Validate security controls and audit logging

Phase 2: Production Deployment (Weeks 3-4)

- Connect to production database with read-only credentials
- Configure enterprise authentication (LDAP/SSO integration optional)
- Load knowledge base documents for RAG functionality
- Establish monitoring and alerting infrastructure

Phase 3: Enterprise Rollout (Weeks 5-8)

- Onboard department champions and power users

- Conduct training sessions (minimal—1 hour per user)
- Expand to additional databases and data sources
- Implement advanced features (custom embeddings, fine-tuned prompts)

Deployment Options

Cloud Deployment

Deploy on AWS, Azure, or GCP using managed Kubernetes (EKS/AKS/GKE) or container services (ECS/ACI/Cloud Run).

On-Premises

Air-gapped deployment using local LLM inference with llama.cpp. Ideal for highly regulated industries.

Hybrid Architecture

Database on-premises with LLM API in cloud, or vice versa. Flexibility to meet security and compliance requirements.

Multi-Tenant SaaS

Managed service option with data isolation, automatic updates, and 24/7 support (roadmap item).

Integration & Extensibility

CGI integrates seamlessly with existing enterprise infrastructure:

- **Database Connectivity:** PostgreSQL (native), MySQL/MariaDB, SQL Server, Oracle (adapters)
- **Authentication:** LDAP, Active Directory, SAML 2.0, OAuth 2.0
- **Monitoring:** Prometheus metrics, Grafana dashboards, CloudWatch/Azure Monitor integration
- **API Layer:** RESTful API for embedding CGI in existing applications

Investment & ROI

Licensing Model

CGI is offered as an **open-source solution** with commercial support and enterprise features available. This proof-of-concept demonstrates the core capabilities deployable in your environment today.

$0
Open Source Core

2-4
Weeks to Production

60-80%
Query Workload Reduction

Expected ROI

For a mid-sized organization (500-1000 employees) with a 5-person data analytics team:

- **Analyst Time Savings:** 2-3 FTE reclaimed for strategic work (~$200-300K/year value)
- **Decision Velocity:** 10x faster data access enables faster pivots and market response
- **Risk Avoidance:** Prevent costly database incidents from user error (average: $500K-$2M per incident)
- **Productivity Gains:** 100+ business users gain self-service data access (conservatively $50K/year in productivity)

Estimated Annual Value: $300K - $500K for a typical mid-market deployment

Transform Your Data Access Strategy

CGI delivers immediate value with minimal investment. Deploy a pilot in your environment, validate ROI with real users and real data, then scale enterprise-wide.

Next Steps

1. Proof of Concept (Week 1)

Deploy CGI on a non-production database with representative data. Invite 5-10 users to evaluate the interface and provide feedback.

2. Pilot Program (Weeks 2-4)

Expand to production database with read-only credentials. Onboard 25-50 users from key departments. Measure query reduction on analytics team and user satisfaction scores.

3. Enterprise Rollout (Weeks 5-8)

Scale to organization-wide deployment. Integrate with authentication systems, expand to additional databases, implement governance policies.

4. Continuous Optimization (Ongoing)

Refine prompts based on usage patterns, expand knowledge base, add custom integrations, monitor security and compliance.

Why Act Now

- **Competitive Advantage:** Early adopters gain 6-12 month lead in data democratization
- **Low Risk:** Open-source core with no vendor lock-in; deploy and evaluate at no cost
- **Proven Technology:** Built on enterprise-grade PostgreSQL, Docker, and LLM infrastructure
- **Rapid Value:** See results in weeks, not quarters

CGI - Conversational Database Interface
Security-First Data Democratization for the Enterprise

For more information or to schedule a demonstration, contact your solution architect.