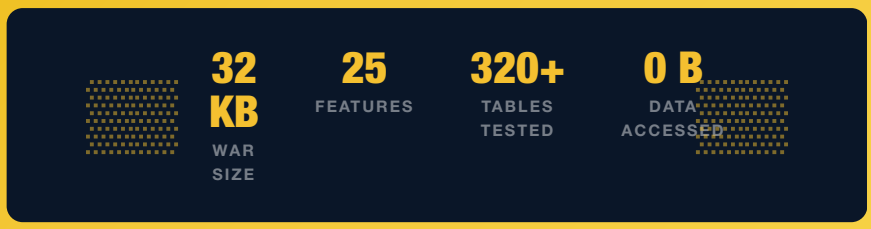


**Cartabase v2.0 â€” Deployed as a single WAR file on IBM WebSphere Liberty**

# Cartabase

Cartabase Enterprise Schema Intelligence Platform for IBM DB2 is purpose-built for air-gapped, distributed defence and government environments. Deployed as a single 32 KB WAR file on IBM WebSphere Liberty, Cartabase reads only structural metadata from DB2 system catalog views â€” never application data â€” and delivers 25 analytical capabilities including automatic foreign-key inference, multi-level impact analysis, schema fingerprinting, and NL-to-SQL conversion. The platform runs on IBM Power P10/P11 (PPC64LE) infrastructure with zero outbound network connections.

## Highlights

- World's smallest enterprise schema tool â€” 32 KB total footprint, 1000x lighter than alternatives
- Zero-data-access architecture â€” reads ONLY SYSCAT system catalog views, never application tables
- Fully air-gap compatible with zero outbound network connections â€” verified through traffic analysis
- Automatic FK inference discovers relationships without declared constraints â€” unique capability
- SHA-256 schema fingerprinting for cross-environment consistency verification across 1,000+ servers
- NL-to-SQL conversion without AI/LLM â€” fully deterministic, works identically in disconnected environments
- Multi-level impact analysis with 4-tier risk scoring before any schema change
- Compliant with DPDP Act 2023, IT Act 2000, MeitY, Defence CDA/MoD, SOX, GDPR, ISO 27001
- Flat-rate perpetual licensing with no per-seat or recurring costs

Cartabase enables database administrators to see and understand their entire DB2 schema landscape â€” every table relationship, every dependency chain, every structural risk â€” without ever touching operational data. The controlling principle of Cartabase is zero-data-access: the DBA runs a single read-only query against DB2 system catalog views (SYSCAT.TABLES, SYSCAT.COLUMNS, SYSCAT.REFERENCES, SYSCAT.INDEXES), exports the result as a CSV, and uploads it through the web interface. The uploaded metadata â€” table names, column names, data types, and foreign key definitions â€” is all Cartabase ever sees. Application tables and business data are architecturally unreachable, enforced by design rather than by access controls.

## Key Schema Intelligence Features

### Discovers Table Relationships Without Declared Foreign Key Constraints

- Analyzes column naming conventions, data types, and structural patterns to infer foreign-key relationships â€” critical in legacy DB2 environments where constraints were never declared or were dropped for performance.
- Assigns confidence scores (High, Medium, Low) to each discovered relationship, allowing DBAs to prioritize validation and review efforts.
- Tested on a 320-table enterprise schema, correctly inferring 103 foreign-key relationships spanning HR, Finance, and Operations domains.
- Works identically across all database sizes â€” from 12-table departmental databases to 320+ table enterprise schemas.

### Predicts the Full Downstream Impact of Schema Changes Before Execution

- Traverses the relationship graph to identify every table affected by a proposed structural modification â€” including indirect dependencies through intermediate tables at unlimited cascade depth.
- Each affected table receives a risk classification (CRITICAL, HIGH, MEDIUM, LOW) based on its position in the dependency chain, the confidence of connecting relationships, and its fan-out degree.
- Impact analysis completes in under 2 seconds regardless of schema size, with results showing hop distance and the exact relationship path through which changes propagate.
- Auto-generates the SQL JOIN chain connecting the source table to each affected downstream table.

### Verifies Schema Identity Across Distributed Environments with Cryptographic Fingerprinting

- Generates a deterministic SHA-256 fingerprint for every uploaded schema that is independent of CSV row ordering â€” two databases with identical structures produce identical fingerprints regardless of when or how the CSV was exported.
- Schema Drift Detection compares two fingerprints and reports every structural difference: added tables, removed columns, type changes, and index modifications.
- Enables fleet-wide schema consistency verification across hundreds of distributed servers â€” a single mismatched fingerprint identifies unauthorized structural modifications.
- Comparison completes in under 3 seconds with a detailed diff report of all structural changes.

### Converts Natural Language to SQL JOINs Without AI, LLM, or Cloud Dependency

- Translates plain English queries into syntactically correct multi-table SQL JOINs by traversing the schema graph and finding the optimal confidence-weighted path between referenced tables.
- The conversion is fully deterministic â€” no machine learning model, no large language model, no cloud API call. Works identically in air-gapped environments.

• Auto-generates ready-to-execute SQL with proper table aliases, JOIN conditions, and WHERE clauses.

## Provides Objective Schema Quality Assessment with Graph-Theoretic Metrics

- Computes 8 graph-theoretic metrics for every table: cyclomatic complexity, Shannon entropy, coupling/cohesion ratios, fan-in/fan-out degree, and betweenness centrality.
- Generates an overall schema health score (0-100) that identifies specific violations â€" orphaned tables, over-coupled hubs, normalization failures â€" giving DBAs an objective quality assessment.
- Performs per-table normalization analysis (1NF/2NF/3NF) with specific violation identification and remediation guidance.
- Topology signature combines SHA-256 hash with betweenness centrality ranking, providing a structural identity beyond simple column comparison.

## Generates Interactive Visual ERD with Zero External Dependencies

- Renders an interactive force-directed graph with color-coded domain clusters â€" built entirely in vanilla JavaScript with zero external library dependencies.
- Supports pan, zoom, and click-to-inspect on schemas up to 320+ tables with layout auto-stabilization in under 3 seconds.
- Nodes are colored by domain cluster (detected via Union-Find algorithm), allowing visual identification of schema subsystems at a glance.

## Produces Compliance-Ready Documentation for Regulated Environments

- **Data Dictionary:** Generates an audit-ready table and column catalog with data types, nullability, and relationship documentation â€" formatted for direct submission to compliance auditors.
- **Complete Audit Trail:** Logs every operation with timestamp, user identity, action performed, and source IP address. No action can be taken without an audit record.
- **Authentication & RBAC:** Built-in role-based access control with session management. All pages require authentication; unauthenticated requests redirect to login.
- Meets compliance requirements for DPDP Act 2023, IT Act 2000, MeitY Guidelines, Defence CDA/MoD, SOX, GDPR, and ISO 27001.

## Deploys in Under 2 Minutes with Zero Configuration

- Single 32 KB WAR file dropped into the IBM WebSphere Liberty `dropins/` directory â€" server starts in under 5 seconds.
- No installation wizard, no configuration file, no database connection, no network setup.
- Pre-built container images available for x86_64 and PPC64LE (IBM Power P10/P11) architectures.
- Optional central consolidation: 1 server + 500 GB for fleet-wide schema monitoring across all establishments.
- Offline license validation via SHA-256 signed keys â€" no call-home, no activation server.

## Enables Cross-Schema Analysis and Path Discovery

- **Optimal Path Finder:** Discovers confidence-weighted JOIN paths between any two tables with weakest-link analysis, automatically generating the SQL.
- **Common Ancestor Detection:** Identifies shared parent tables between any pair of tables in the schema.
- **Cross-Schema Correlation:** Matches semantically equivalent columns across different databases with similarity scoring (e.g., ITEM_ID in one schema mapping to PART_NO in another).

- **Data Lineage:** Traces the full upstream dependency chain for any table, showing every table it transitively depends on.

| Table 1: Cartabase Performance and Capacities | | | |
|---|---|---|---|
| | **Small Schema (10-50 tables)** | **Medium Schema (50-200 tables)** | **Large Schema (200-500 tables)** |
| **Schema mapping time** | < 2 seconds | < 10 seconds | < 30 seconds |
| **FK inference** | All patterns detected | All patterns detected | All patterns detected |
| **Impact analysis response** | < 500 ms | < 1 second | < 2 seconds |
| **Schema fingerprint generation** | < 200 ms | < 500 ms | < 1 second |
| **Schema drift comparison** | < 1 second | < 2 seconds | < 3 seconds |
| **Visual ERD render** | < 1 second | < 2 seconds | < 3 seconds |
| **JVM heap usage** | < 16 MB | < 32 MB | < 64 MB |

Note: Results measured on IBM DB2 12.1.5.0. Schema mapping time includes CSV parsing, relationship inference, and graph construction. Impact analysis measured with 5+ cascade levels enabled.

| Table 2: Cartabase Analytical Capabilities | |
|---|---|
| **Schema Discovery & Mapping** | |
| **Automatic FK inference** | Pattern-based, confidence-scored (High/Medium/Low), no declared constraints required |
| **Visual ERD** | Interactive force-directed graph, color-coded domain clusters, zero external libraries |
| **Relationship map** | Full FK listing with confidence levels and relationship type |
| **Cluster detection** | Union-Find algorithm for automatic domain/subsystem identification |
| **Column type distribution** | Data type frequency analysis and visualization |
| **Impact & Risk Analysis** | |
| **Impact analysis** | Multi-level cascade prediction, unlimited depth, 4-tier risk scoring (CRITICAL/HIGH/MEDIUM/LOW) |
| **Risk scoring** | Per-table risk classification based on fan-out, centrality, and cascade position |
| **Data lineage** | Full upstream dependency chain tracing |
| **Path Discovery & SQL Generation** | |
| **Find path** | Shortest path between any two tables with auto-generated SQL JOIN |
| **Optimal path** | Confidence-weighted path with weakest-link analysis |
| **Common ancestors** | Shared parent table identification for any table pair |
| **NL-to-SQL** | |

| | |
|---|---|
| | Natural language to SQL JOIN conversion â€” deterministic, no AI/LLM |
| **Join builder** | Multi-table JOIN generation from table name selection |
| **Schema Quality & Integrity** | |
| **Schema fingerprint** | SHA-256, deterministic, order-independent hash for identity verification |
| **Schema drift detection** | Structural comparison: added/removed/modified tables, columns, indexes |
| **Health score** | 0-100 score with specific violation identification |
| **Normalization analysis** | Per-table 1NF/2NF/3NF assessment with violation details |
| **Complexity metrics** | Cyclomatic complexity, Shannon entropy, coupling, cohesion, fan-in/out, centrality |
| **Topology signature** | SHA-256 structural hash + betweenness centrality ranking |
| **Cross-schema correlation** | Semantic column matching across databases with similarity scoring |
| **Compliance & Administration** | |
| **Data dictionary** | Audit-ready compliance documentation, auto-generated from schema metadata |
| **Audit log** | Timestamp, user, action, IP address for every operation |
| **Authentication & RBAC** | Built-in role-based access control with session management |

| Table 3: Cartabase Platform Specifications | |
|---|---|
| **Application Server** | |
| **Required runtime** | IBM WebSphere Liberty 23.x or later |
| **Java runtime** | Java SE 8, 11, 17, or 21 |
| **Package format** | WAR (Web Application Archive) â€” single file, 32 KB |
| **Deployment method** | Drop WAR into Liberty `dropins/` directory â€” no configuration required |
| **Database Compatibility** | |
| **Supported databases** | IBM DB2 10.5, 11.1, 11.5, 12.1 |
| **Data source** | Pipe-delimited CSV exported from SYSCAT system catalog views |
| **Catalog views used** | SYSCAT.TABLES, SYSCAT.COLUMNS, SYSCAT.REFERENCES, SYSCAT.INDEXES |
| **Database connection** | None â€” Cartabase never connects to any database |
| **Supported Architectures** | |
| **x86_64** | Linux, Windows, AIX â€” any OS with Java SE support |
| **PPC64LE** | IBM Power P10, P11 â€” pre-built container image available |
| **Container support** | Docker 20.x+, Podman 4.x+, OpenShift 4.x+ |
| **Resource Requirements** | |
| **Disk footprint** | < 1 MB total (runtime) |
| **Memory** | < 64 MB JVM heap (320-table schema) |
| **CPU** | Single core sufficient â€” no background threads or workers |
| **Network** | Zero outbound connections â€” fully air-gap compatible |

| Table 4: Security and Compliance | |
|---|---|
| **Data Access** | |
| **Business data access** | Zero â€” architecturally enforced by design, not by access controls |
| **Data in transit** | CSV uploaded via browser â€” HTTPS when Liberty configured with TLS |
| **Data at rest** | Schema metadata exists only in JVM memory during session â€” no persistent storage |
| **Network posture** | Zero outbound connections â€” no DNS, HTTP, HTTPS, or any network calls |
| **Authentication & Authorization** | |
| **Authentication** | Built-in login with session management |

| | |
|---|---|
| **Authorization** | Role-based access control (RBAC) |
| **Audit trail** | Timestamp, user, action, IP address logged for every operation |
| **Input sanitization** | All input HTML-escaped before rendering â€” XSS protection |
| **License & Validation** | |
| **License model** | Perpetual site license â€” flat-rate, no per-seat or recurring fees |
| **License validation** | SHA-256 offline validation â€” no activation server, no call-home |
| **Trial availability** | 60-day trial key included |
| **Regulatory Compliance** | |
| **Indian standards** | DPDP Act 2023, IT Act 2000, MeitY Guidelines, Defence CDA/ MoD |
| **International standards** | SOX, GDPR, ISO 27001 |

| Table 5: Deployment Options | |
|---|---|
| **Standalone** | |
| **WAR on Liberty** | Drop `xray.war` (32 KB) into Liberty `dropins/` â€” starts in < 5 seconds |
| **Containerized** | |
| **Docker (x86_64)** | Pre-built image, < 1 MB application layer |
| **Docker (PPC64LE)** | Pre-built image for IBM Power P10/P11 |
| **OpenShift** | Deployment manifest available, health probe on `/health` |
| **Central Consolidation (Optional)** | |
| **Central server** | 1 server + 500 GB for fleet-wide schema view across all establishments |
| **Schema collection** | CSVs from distributed sites uploaded to central instance for comparison |

| Table 6: Ordering Information | |
|---|---|
| **CB-LIC-SITE** | Cartabase Site License â€” Perpetual, covers all servers at one establishment |
| **CB-LIC-FLEET** | Cartabase Fleet License â€” Perpetual, covers all servers across all establishments |
| **CB-SUP-5Y** | 5-Year Support & Maintenance â€” includes version upgrades and priority support |
| **CB-IMG-PPC64** | PPC64LE Container Image for IBM Power P10/P11 |
| **CB-CENTRAL** | Central Consolidation Server Setup (1 server + 500 GB storage) |
| **CB-TRAIN** | DBA Training â€” 2 hours per batch, on-site or remote |

cartabase

**RAMS Creative Technologies Pvt. Ltd.**
1st Floor, C-61 A, Road No. 1-C, VKI Area
Jaipur, Rajasthan 302016, India

**Sales:** sales@ramscreative.in
**Support:** support@ramscreative.in



cartabase_ds_v2.0_080226