Selvo Lens: Enterprise Data Intelligence with Verified Computations

A White Paper on Execution-Verified Analytics for Business-Critical Data

January 2026 | Selvo Lab

Executive Summary

Enterprise organizations face a paradox: artificial intelligence systems that can answer natural language questions about business data have become indispensable, yet their fundamental unreliability makes them unsuitable for decision-making in regulated industries, financial analysis, or any domain where accuracy matters.

The industry standard - Retrieval-Augmented Generation (RAG) - attempts to solve this by grounding LLM responses in source documents. However, this approach merely reduces hallucination; it does not eliminate it. The LLM still "synthesizes" answers, introducing uncertainty into every response.

Selvo Lens introduces a fundamentally different paradigm: rather than asking LLMs to reason about data, we use them to generate executable code that computes answers. The result is execution-verified computation with auditability.

Key Claims:

- Computed results are execution-verified (numbers are computed, not invented)
- Source grounding for document answers + audit logging for traceability
- Transparent computations (code/methodology available, configurable)
- Designed to support regulated workflows (auditability + self-hosted deployment)
- Self-hosted deployment (on-premises or in your private cloud account)

The Business Problem

Why Current AI Systems Fail in Enterprise

Many enterprises have experimented with large language models for business intelligence, but struggle to deploy them for business-critical decisions. A consistent blocker is **trust**.

Industry Data (estimates from public research):

- Hallucination risk varies widely by model, prompt, and task complexity.
- Enterprise adoption barrier: "We cannot rely on answers we cannot verify" (common feedback in regulated environments)
- Compliance teams block AI implementation in regulated environments.
- Finance departments mandate human review for all AI-generated insights.

The RAG Approach Doesn't Solve This

Traditional RAG (Retrieval-Augmented Generation) improves on base LLMs by adding source documents to the prompt. This reduces hallucination but does not eliminate it:

Where RAG Still Fails:

1. The LLM still interprets retrieved data subjectively

2. Complex multi-step reasoning requires chain-of-thought prompting (which adds error)
3. Calculated results (sums, averages, percentages) rely on LLM arithmetic
4. No way to audit "how" the AI reached its conclusion beyond human reading
5. Regulatory compliance officers remain skeptical

Real Example:

- User asks: "What is our Q3 revenue by region?"
- RAG system retrieves relevant documents
- LLM synthesizes: "Based on the documents provided, Q3 revenue appears to be approximately $X"
- Problem: an approximate answer is unacceptable in finance.

The Selvo Lens Solution

Core Innovation: Computed Answers, Not Generated Answers

Rather than asking the LLM to synthesize an answer, Selvo Lens uses the LLM to generate executable code that computes the answer against your actual data.

The Computation Model:

1. User asks a question in natural language
2. Selvo Lens extracts schema and context from your data
3. LLM generates precise, executable code (pandas/SQL)
4. Code is validated and executed in a sandboxed environment
5. Result is computed via executed code and traceable

Real Example (Same Question):

- User asks: "What is our Q3 revenue by region?"
- System generates: `df[df['Quarter'] == 'Q3'].groupby('Region')['Revenue'].sum()`
- System executes on actual data
- System returns: an exact breakdown computed from your data
- Difference: Exact answer with auditability

Why This Approach Guarantees Accuracy

The accuracy guarantee applies to **code execution**, not to LLM reasoning:

1. **Mathematical Verification**: When code executes successfully, the computation is mathematically correct
2. **Schema-Aware Generation**: LLM knows exact column names and types, preventing hallucinated fields
3. **Self-Correction**: If code fails, system automatically corrects based on error message (up to 3 retries)
4. **Security Validation**: No file access, no external calls, no dangerous functions
5. **Audit Trail**: Responses can include generated code/methodology (configurable)
6. **Deterministic Execution**: Same code + same data = same result; low-temperature settings reduce variance

Important: We distinguish between two types of accuracy:

- **Computed Output Correctness**: When code executes, returned numbers are mathematically correct for that computation.

- **Intent Recognition**: Whether the system understood what the user asked for (varies by dataset, schema quality, and query phrasing).

If the system misinterprets the question, the user can rephrase the request and rerun.

Key Insight: We don't trust the LLM to reason correctly. We only trust the LLM to generate reasonable code that we can validate, execute, and verify. The LLM is a **code generator**, not an analyst.

Technical Architecture

System Overview

Data Flow:

1. **User** sends natural language question via browser
2. **Selvo Lens Engine** extracts schema from your data
3. **LLM** generates executable code based on schema
4. **Validation Layer** checks code for security and correctness
5. **Execution Sandbox** runs code against actual data
6. **User** receives computed answer + audit logging, with optional code/methodology (configurable)

Components:

- API Server (FastAPI)
- LLM Engine (Ollama - runs locally)
- Vector Database (ChromaDB)
- Code Execution Sandbox

Supported Data Sources:

- Excel, CSV
- PDF, Word documents
- Structured exports from databases (CSV/Excel)
- Images with OCR

System Design

Selvo Lens consists of five integrated layers:

Layer 1: Intelligent Query Routing

- Classifies incoming queries as document Q&A, analytics query, or hybrid
- Routes to appropriate processing pipeline
- Ensures optimal accuracy for each query type

Layer 2: Schema-Aware Code Generation - Dynamically extracts column names, data types, and sample values from your data. - Injects schema into LLM prompt to prevent hallucinated column names. - Uses Chain-of-Thought prompting: "think step by step" - Includes few-shot examples of similar queries - Generates precise, executable code with low temperature (0.1)

Layer 3: Multi-Stage Validation - Syntax checking: Is the code valid Python/SQL? - Security scanning: Forbidden patterns (imports, file operations, exec, eval) - Schema verification: Do referenced columns actually exist? - Type checking: Are operations compatible with data types?

Layer 4: Safe Execution Engine - Sandboxed Python environment (no file access, no network) - Pandas/NumPy computation on actual DataFrame - Self-correction loop: up to 3 retries with error feedback - Automatic retry if code fails

Layer 5: Verified Response

- Returns computed results for analytics, and grounded answers for document Q&A.
- Optionally includes generated code and/or methodology (configurable).
- Includes record counts for applicable queries.
- Creates an audit trail for traceability.

Accuracy Guarantees

What "Verified" Means

We must be precise about what we claim:

Metric	Definition	How Selvo Lens Addresses It
Computed Output Correctness	Are returned numbers mathematically correct?	Computed by executed code (not LLM arithmetic)
Intent Recognition	Did system understand what you asked?	Measured during pilot; improves with schema quality + prompt tuning
Query Success Rate	Did system return a result (vs. error)?	Uses validation + up to 3 self-correction retries
Safety Constraints	Can generated code escape the sandbox?	Validated + sandboxed execution (no file/network access)

When Selvo Lens returns a computed number from executed code, that number is mathematically correct for that computation because it was computed from your data, not generated by an LLM.

Comparison to Industry Alternatives

Aspect	Standard LLM Chat	Standard RAG	**Selvo Lens**
Numeric analytics	Model-generated text/arithmetic	Model-generated synthesis	Execution-verified computations
Verifiability	Limited	Partial (sources)	Computations reproducible; sources for document answers
Auditability	Low	Medium	High (logs + reproducible steps)

Aspect	Standard LLM Chat	Standard RAG	**Selvo Lens**
Code/methodology visibility	Rare	Rare	Configurable (can be shown or exported)

Enterprise Applications

Use Case 1: Financial Analysis

Scenario: CFO needs Q3 revenue breakdown by product line and customer segment

Traditional Approach: - Finance analyst manually pulls data from multiple systems - Creates pivot table in Excel - Shares with CFO for review - Turnaround time: 2-4 hours

Selvo Lens Approach: - CFO: “Show me Q3 revenue by product line and segment” - System generates: `df[df['Quarter'] == 'Q3'].groupby(['Product_Line', 'Segment'])['Revenue'].sum().sort_values(ascend` - System executes and returns exact breakdown - CFO sees results with audit logging for traceability - Turnaround time: seconds to minutes (depends on data size and hardware)

Value: Eliminates manual analysis errors, enables ad-hoc exploration, audit-ready results

Use Case 2: Compliance Reporting

Scenario: Compliance officer must report on data retention policy violations

Traditional Approach: - Data analyst queries database manually - Creates compliance report in Word - Report is reviewed and signed off - Regulators ask: “How did you calculate this?” - Answer: “Analysis team reviewed the data”

Selvo Lens Approach: - Officer: “Show me all customer records older than 7 years” - System generates: `df[df['RecordAge'] > 7]['CustomerId'].count()` - System returns exact count (and can include the code/methodology that generated it) - Officer includes both result AND code in regulatory submission - Regulator can verify: “Yes, this code correctly counts old records”

Value: Regulatory defensibility, faster compliance cycles, reduced audit friction

Use Case 3: Sales Intelligence

Scenario: Sales VP needs top customers by region for Q4 planning

Selvo Lens Approach: - VP: “Top 10 customers by revenue in EMEA region” - System generates and executes code - Returns exact top 10 with verifiable computations - Can immediately drill down: “Show me pipeline for customer XYZ” - All answers backed by code, not assumptions

Value: Trust in data, faster decision-making, confidence in strategy

Use Case 4: HR Analytics

Scenario: HR Director needs diversity metrics for reporting

Selvo Lens Advantage: - All metrics are computed from actual employee data. - Gender distribution: Exact count with breakdown by department, level, tenure. - Pay equity analysis: Exact calculations with methodology available. - Regulatory reporting: Defensible because computation is transparent.

Value: Confidence in HR metrics, regulatory compliance, no disputes about "how we calculated this"

Implementation & Deployment

Data Sources Supported

- **Structured Data**: Excel, CSV (and structured exports from databases)
- **Documents**: PDF, Word, Markdown, plain text
- **Scanned Documents**: OCR support for scanned PDFs and images (PNG, JPG, TIFF, BMP)
- **Multiple Formats**: Hybrid analysis across different data types

Self-Hosted Deployment (On-Prem or Private Cloud)

Selvo Lens is designed to be **self-hosted**: deploy it on-premises or inside your private cloud account (for example, dedicated AWS EC2 instances in your VPC).

Why Self-Hosted?

The core value proposition of Selvo Lens is verifiable computation plus strong data control:

1. **Data control**: Keep data inside your infrastructure and security boundary
2. **Residency & governance**: Supports internal data residency and governance requirements
3. **Auditability**: Keeps execution, logs, and artifacts within your environment
4. **Cost predictability**: Avoids per-request SaaS pricing; infrastructure costs are transparent

Deployment Architecture

Standard Self-Hosted Deployment:

```
Your Infrastructure
├── Data Sources (Files; structured exports from databases)
├── Selvo Lens Engine (Docker containers)
│   ├── API Server
│   ├── LLM Engine (Ollama - local)
│   ├── Code Execution Sandbox
│   └── Security Layer
└── Authentication & Audit Logging
```

Typical Timeline:

- Assessment & planning: 3-5 days
- Infrastructure setup: 1-2 weeks
- Integration with data sources: 1-2 weeks
- Pilot with your data: 2-4 weeks
- Full deployment: 6-8 weeks total

System Requirements

Minimum Configuration:

- 32GB RAM (24GB reserved for LLM)
- 8 CPU cores

- 100GB SSD storage
- CPU-only supported (slower); NVIDIA GPU with 24GB VRAM recommended (RTX 3090, RTX 4090)
- Docker & Docker Compose support
- Direct access to data sources (file systems and/or database exports)

Production Configuration (Recommended):

- 64GB+ RAM
- 16+ CPU cores
- NVIDIA RTX 3090/4090 (24GB VRAM)
- 500GB+ NVMe SSD storage
- Docker Compose orchestration (no Kubernetes needed)
- Network isolation
- Backup and disaster recovery

Infrastructure Deployment:

- On-premises: Bare metal servers or hypervisor (ESXi, Hyper-V)
- AWS private deployment: EC2 instances in private VPC with VPN

Why Not Kubernetes?

Kubernetes adds operational complexity without addressing Selvo Lens requirements: - Single-instance deployment model (one container per organization) - No horizontal scaling needed (data lives on-premises) - Docker Compose provides all necessary orchestration. - Kubernetes expertise is expensive and unnecessary for this use case.

Standard deployment: **Single Docker Compose stack per organization**, scaled to your workload and availability needs.

ROI Analysis

Conservative ROI Estimate

Important Note: ROI is organization-specific and depends on current data workflows. The following estimates are conservative and apply to organizations that currently rely on manual analysis for business-critical data.

Example: Mid-size Financial Services Firm (100+ employees, $500M+ AUM)

Annual Cost Reduction

Cost Category	Current State	With Selvo Lens	Annual Savings
Analyst time on ad-hoc queries	5 analysts × 1,000 hrs/yr × $75/hr = $375K	Same analysts, 70% time saved on data tasks	**$112.5K**
Compliance report preparation	Manual verification, 500 hrs/yr × $100/hr = $50K	Automated with audit trail, 80% time reduction	**$40K**
Report review/sign-off time	Finance director reviews every report (200 hrs/yr)	Director only spot-checks (75% reduction)	**$15K**
Audit preparation costs	Reconstruct methodology (40 hrs × $150/hr = $6K)	Audit logging and reproducible steps available	**$6K**
Regulatory penalties (avoided)	Risk of calculation errors/year × potential penalty	Reduced risk via reproducible computations and review	**Varies**

Conservative Annual ROI: $190-200K Selvo Lens subscription: $4,999/month (~$60K/year, includes implementation and support) **Payback period: 3-4 months**

Risk Reduction Value (Unquantified)

1. **Regulatory Defense**: Audit-ready calculations reduce regulatory scrutiny and defense costs
2. **Compliance Risk**: Eliminates reputational damage from calculation errors in reports
3. **Decision Confidence**: Finance leadership makes decisions on verified data, not estimates
4. **Audit Cycle Time**: External auditors accept methodology faster (code is transparent)

What This ROI Assumes

- Organization has 50+ employees doing analytical work
- Analysts spend 20-30% of time on manual data extraction/verification
- Compliance reporting requires significant manual labor
- Organization is in regulated industry (finance, healthcare, legal)
- Current error rate on manual analysis is 2-5%

What This ROI Does NOT Account For

- Potential gains from faster decision-making
- Elimination of disputes over “how was this calculated?”
- Reduced training time for new analysts (code is self-documenting)
- Value of audit trail for internal investigations
- Competitive advantage from faster reporting

Conservative Estimate: $190-200K annual savings Realistic Value (including intangibles): $300-500K+

Security & Governance

Data Sovereignty

- Self-hosted deployment (on-premises or private cloud)
- Designed for operation without third-party data processing
- Local LLM execution (Ollama)
- Complete control over access and retention

Security Features

- JWT-based authentication
- Role-based access support (JWT auth + roles)
- Tenant isolation support when enterprise configuration is enabled
- Prompt injection defenses (context delimiting + detection/logging)
- Sandboxed code execution
- Audit logging of all queries and results

Compliance

Selvo Lens architecture is designed to support compliance requirements in regulated industries:

- **GDPR**: Supports data residency and internal governance controls when self-hosted.
- **HIPAA**: Supports auditability and controlled deployment for PHI workflows when combined with appropriate infrastructure controls.
- **SOC 2**: Supports common security-control requirements (auth, audit logs, least privilege) as part of an organization's overall control environment.
- **Financial regulations**: Supports audit trails and reproducible computations for defensible reporting.

Note: Compliance certification is the responsibility of each organization based on their specific implementation and security practices.

Competitive Positioning

Why Selvo Lens Is Different

Capability	ChatGPT	Enterprise RAG	**Selvo Lens**
Can answer any question	Yes	Mostly	Only structured/documented
Verifiable outputs	No	Partial	Execution-verified computations
Regulated workflows	Risky	Risky	Designed for auditability + self-hosting
Deployment model	SaaS	Varies	Self-hosted (on-prem/private cloud)
Code/methodology visibility	No	Rare	Configurable

Capability	ChatGPT	Enterprise RAG	**Selvo Lens**
Audit trail	No	Minimal	Comprehensive
Audit-ready reporting	No	Limited	Designed for defensible reporting
Deterministic execution	No	No	Yes (for executed computations)
Schema-aware generation	No	No	Yes
Self-correction loop	No	No	Yes

The Core Innovation

Other AI systems: LLM generates an answer → You hope it's correct.

Selvo Lens: LLM generates code → Code is validated → Code executes on real data → Result is mathematically correct.

This is not an incremental improvement. It is a fundamentally different approach.

What This Means in Practice

Scenario	ChatGPT/RAG	Selvo Lens
CFO asks for Q3 revenue	"Approximately …"	"Exact total (computed from your data)"
Auditor asks for methodology	"AI analyzed the data"	"Methodology/code available (configurable)"
Regulator requests verification	Cannot provide	Full audit trail with reproducible computation
Same question asked twice	May give different answers	Reproducible when rerun (executed computation)

Why Competitors Cannot Easily Replicate This

1. **Architecture**: Selvo Lens was built from the ground up for code generation and execution - not retrofitted onto a chat interface.
2. **Validation Pipeline**: Multi-layer security scanning, schema verification, and sandboxed execution require deep integration that cannot be added as a feature.
3. **Schema Awareness**: Real-time extraction of column names, data types, and sample values from your actual data - not a static prompt template.
4. **Self-Correction**: Error context is fed back to LLM for automatic retry - requires tight coupling between execution engine and code generation.

The Key Difference: Selvo Lens doesn't ask "Is the LLM right?" Instead, it asks "Is the code correct?" and verifies through execution.

Next Steps

Request a Demo (1-2 hours)

What Happens: 1. You provide sample dataset or connect to data source 2. Our team prepares 3-5 queries from your actual business questions 3. Live demonstration showing: - Natural language question → Code generation - Optional code/methodology review (you can see what will execute) - Live execution on your data - Results with full audit trail - **Self-correction in action**: We intentionally trigger an error and show automatic fix

Expected Outcome: - You see Selvo Lens working on YOUR data. - You understand how results are verified and auditable. - You have a clear picture of implementation requirements.

Preparation Needed (from you): - Sample dataset (5-10K rows, can be anonymized) - 3-5 business questions you currently answer manually - 30 minutes for pre-demo call with our technical team

Pilot Program (2-4 weeks)

Phase 1: Setup (Days 1-3) - Infrastructure assessment and deployment planning - Data source integration and schema analysis - Team training on Selvo Lens query syntax

Phase 2: One Critical Use Case (Days 4-14) - Identify ONE high-impact report or query (e.g., quarterly revenue analysis, compliance data pull) - Deploy Selvo Lens for this specific use case - Compare results to current manual process: - Time to complete report - Accuracy verification - Auditability improvements

Phase 3: Evaluation & Decision (Days 15-21) - Extended testing with 2-3 additional queries - Full team training (data analysts, compliance, finance) - Cost-benefit analysis based on pilot results - Go/no-go decision for full deployment

Success Metrics (Measure During Pilot): - Time reduction: Reduce workflows from hours to minutes where applicable - Accuracy: execution-verified computations vs. current manual error rate - Audit readiness: Can you now provide methodology to auditors? - Team adoption: Are analysts comfortable with the system?

What's Included: - Full technical support from Selvo Lab engineers - Custom configuration for your data formats - Training for 10+ team members - Integration with your existing data sources - Audit logs and documentation - **No licensing commitment** - pilot is fully reversible

Implementation Path (After Pilot Approval)

Week 1-2: Infrastructure Hardening - Production-grade deployment (HA, backup, disaster recovery) - Security review and hardening - Performance optimization (GPU setup if needed)

Week 3-4: Full Rollout - Migration of 10-20 regular reports to Selvo Lens - Extended team training - Monitoring and optimization

Ongoing: - Monthly check-ins for optimization - Support for new use cases and queries - Compliance reporting and audit support

Investment & Timeline

Phase	Duration	Cost	What You Get
Demo	1-2 hours	Free	Proof of concept on your data
Pilot	2-4 weeks	Included	Full evaluation environment
Subscription	Ongoing	$4,999/month	Full system + implementation + support

Compare to Current State: - Analyst cost (5 people): $375K/year - Selvo Lens: $60K/year (all-inclusive) - **Net savings: $130K+ annually** (after Selvo Lens cost)

Contact & Get Started

Initial Consultation (Free, 30 minutes) - Assess your use cases and data landscape - Determine if Selvo Lens is the right fit for your organization - Outline demo and pilot timeline

Contact Selvo Lab: - **Website**: www.selvolab.com - **Email**: contact@selvolab.com - **Response time**: Within 24 hours

What to Prepare for Initial Call: - Description of 2-3 critical queries you currently run manually - Approximate data volume (rows, columns) - Current pain points (accuracy, time, auditability) - Timeline for decision-making

Conclusion

The enterprise AI market has reached an inflection point. Organizations can no longer justify deploying AI systems that generate plausible-sounding answers to business-critical questions. The next generation of enterprise AI must be **verifiable, auditable, and accurate**.

Selvo Lens represents this shift: from "AI that guesses well" to "AI that computes correctly." For any organization that handles regulated data, makes financial decisions, or needs defensible analytics, this distinction is not a nice-to-have - it is mandatory.

The question is no longer "Can AI answer business questions?" It is "Can AI answer them *correctly and verifiably*?"

Selvo Lens answers yes.

Document Version: 1.2 | January 2026