Virtuoso SPIN Inference & Automated Rule Generation

Bridging the gap between complex database logic and business intelligence through scalable, macro-based automation.

PROJECT: CUSTOMER ANALYTICS SPIN RULES | STATUS: COMPLETED (JAN 9, 2026) | STACK: VIRTUOSO 8.0+, W3C SPIN, CLAUDE CODE

Operationalizing Inference at Scale

Transforming manual, error-prone rule creation into a systematic, automated workflow.

01

Analysis (The Pattern)

Codified the tacit knowledge of SPIN rule structure into a documented 6-section pattern derived from reference implementations.

02

Implementation (The Product)

Built '`spin-rule-customer-analytics-1.sql`', implementing 5 dynamic customer segments (Active, At-Risk, High-Value) without data materialization.

03

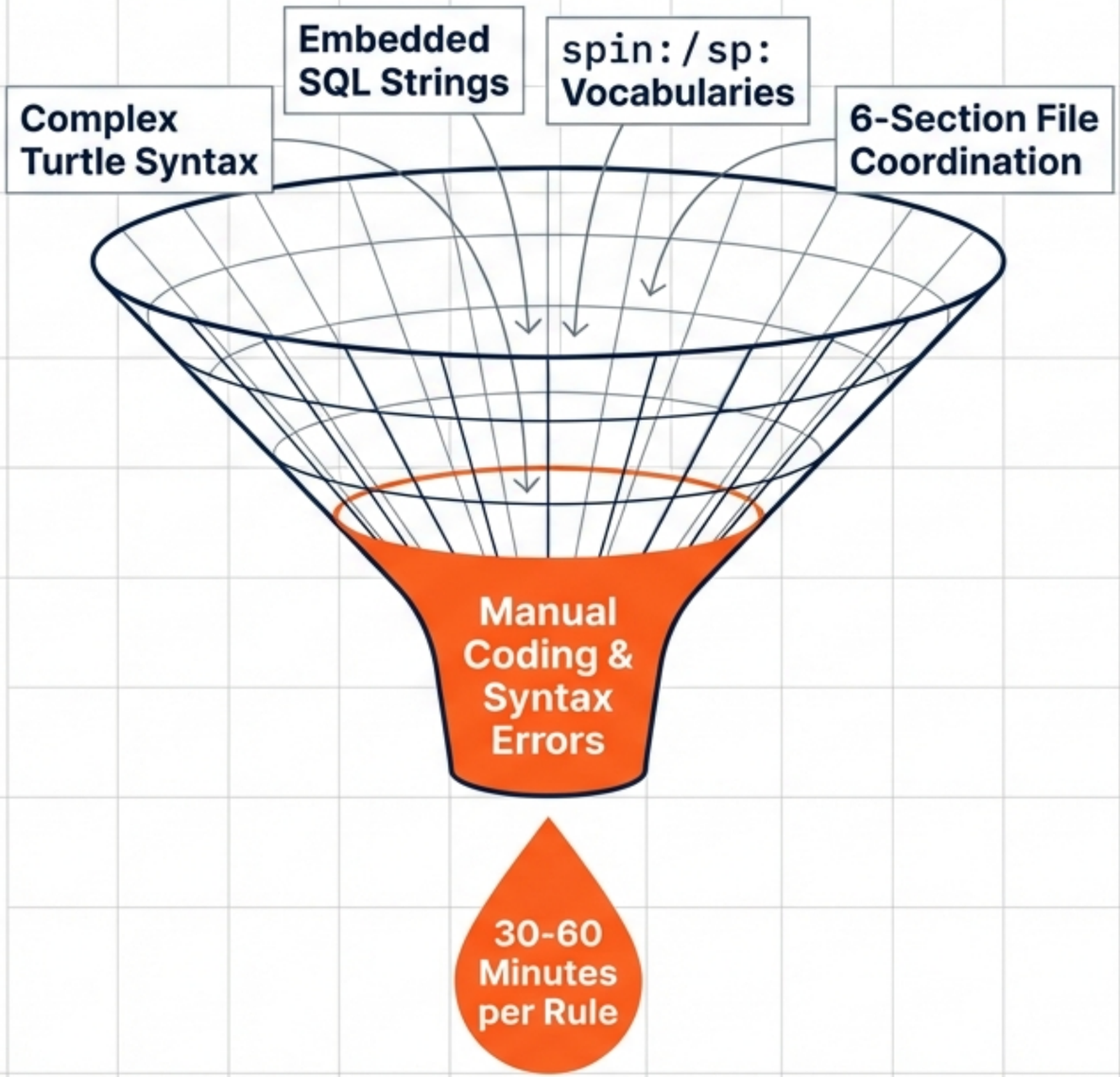
Automation (The Tool)

Developed the '`virtuoso-spin-rule-generator`' skill (11KB), reducing rule creation time from 60 minutes to 2 minutes.

The High Cost of Dynamic Logic

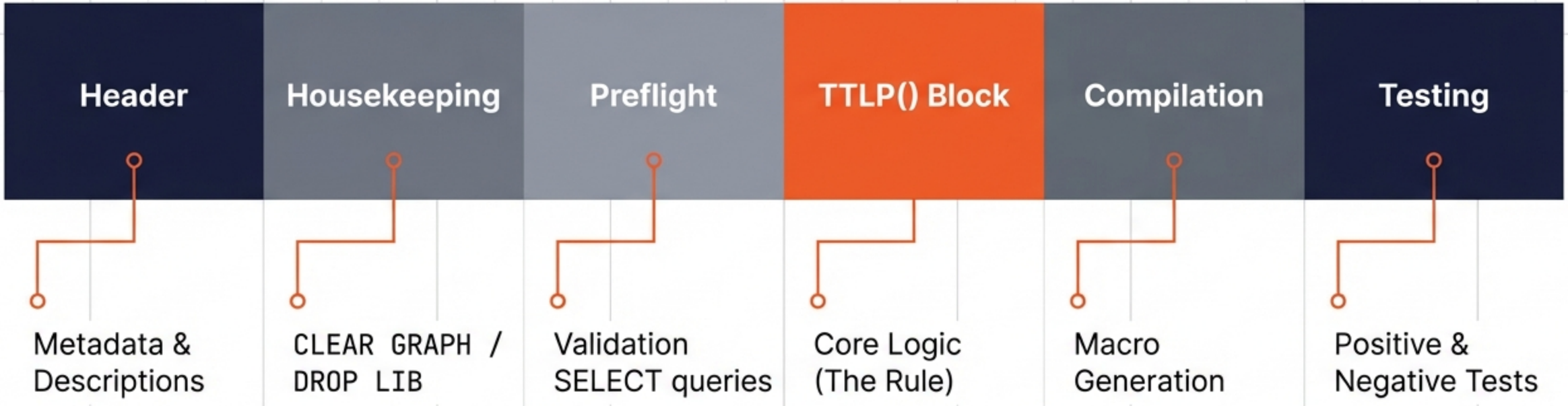
Virtuoso's SPIN offers powerful runtime inference, but manual implementation creates a development bottleneck.

The Friction Funnel



Codifying the 6-Section SPIN Pattern

We identified a rigid, repeatable anatomy for every SPIN rule file to ensure consistency.



Key Insight: Consistency in URN naming is the linchpin of stability.

Proof of Concept: Customer Analytics Module

Classifying customers dynamically based on order recency without ETL.


```
FILE:   spin-rule-customer-analytics-1.sql
LINES:  467
METHOD: MAX(?orderDate) + xsd:duration filters
```

The 5 Dynamic Customer Segments

Class	Time Window	Business Logic	Additional Criteria
`ActiveCustomer`	`< 6 months`	Recent engagement. No intervention needed.	
`AtRiskCustomer`	`6 - 12 months`	Early intervention opportunity.	
`DormantCustomer`	`1 - 2 years`	Re-engagement campaigns required.	
`InactiveCustomer`	`> 2 years`	Win-back strategies.	
`HighValueAtRisk`	`1 - 2 years`	Priority retention effort. “Save this customer”.	`AND 10+ Lifetime Orders`

Anatomy of an Inference Rule

Logic runs at query time via DEFINE input:macro-lib

```
ecrm-ontology:Company
  spin:rule [
    a sp:Construct ;
    sp:text """
       CONSTRUCT { ?company a <#ActiveCustomer> . }
       WHERE {
           {
              SELECT ?company (MAX(?orderDate) AS ?mostRecent)
              WHERE { ... } GROUP BY ?company
           }
           FILTER (?mostRecent > (NOW() - "P6M"^^xsd:duration))
       }
    """
  ] .
```

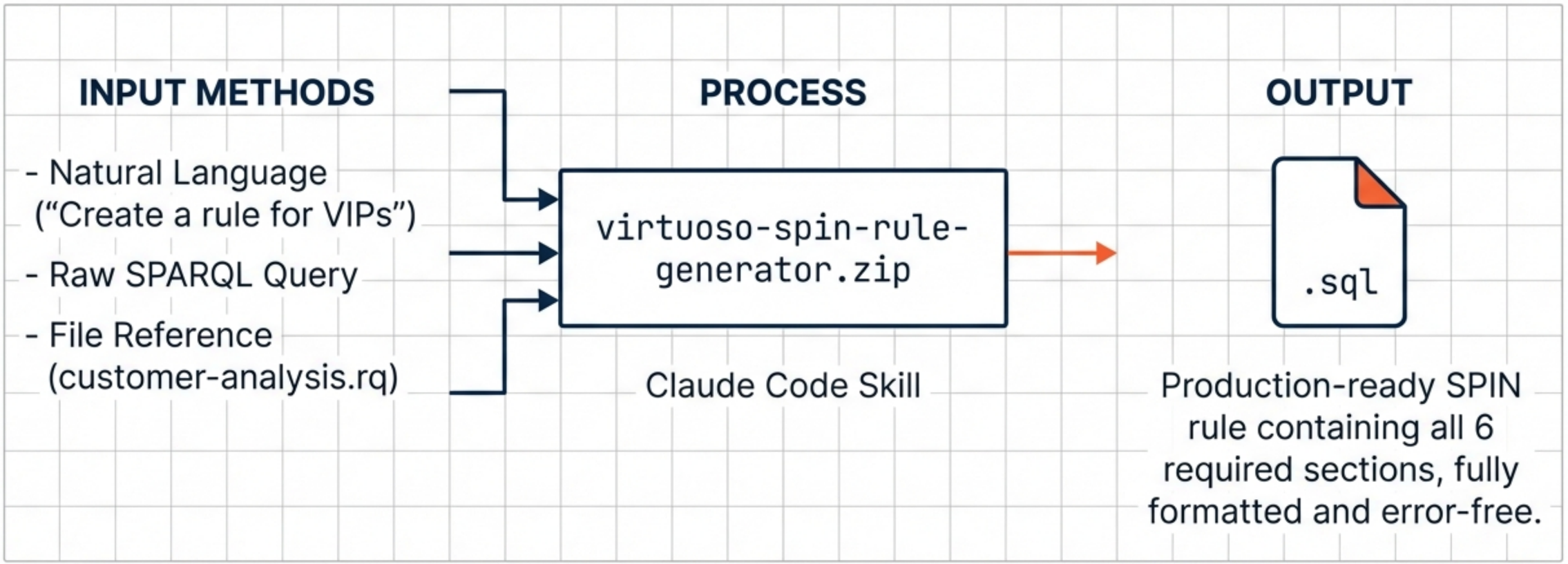
Infers new triples dynamically

Aggregates history

Calculates relative to query time

The Automated Rule Generator Skill

Moving from manual writing to automated generation.



Intelligent Query Transformation

How the generator ensures safety and syntax compliance.

Intent Analysis

Identifies classification criteria and source classes.

Component Extraction

Pulls filters, aggregations, and properties.

Logic Conversion

Translates `SELECT WHERE` logic into `CONSTRUCT WHERE` patterns.

Auto-Testing Generation

Creates `Preflight` validation and **Positive/Negative** verification tests.

Safety Protocol

The skill explicitly handles complex `URN` matching and `xsd:duration` syntax that humans frequently mistake.

Metrics: Manual vs. Automated

MANUAL PROCESS

- 30-60 Minutes / Rule
- High Error Rate (Syntax/URNs)
- Requires Virtuoso Specialist

95% Reduction in Development Time

AUTOMATED SKILL

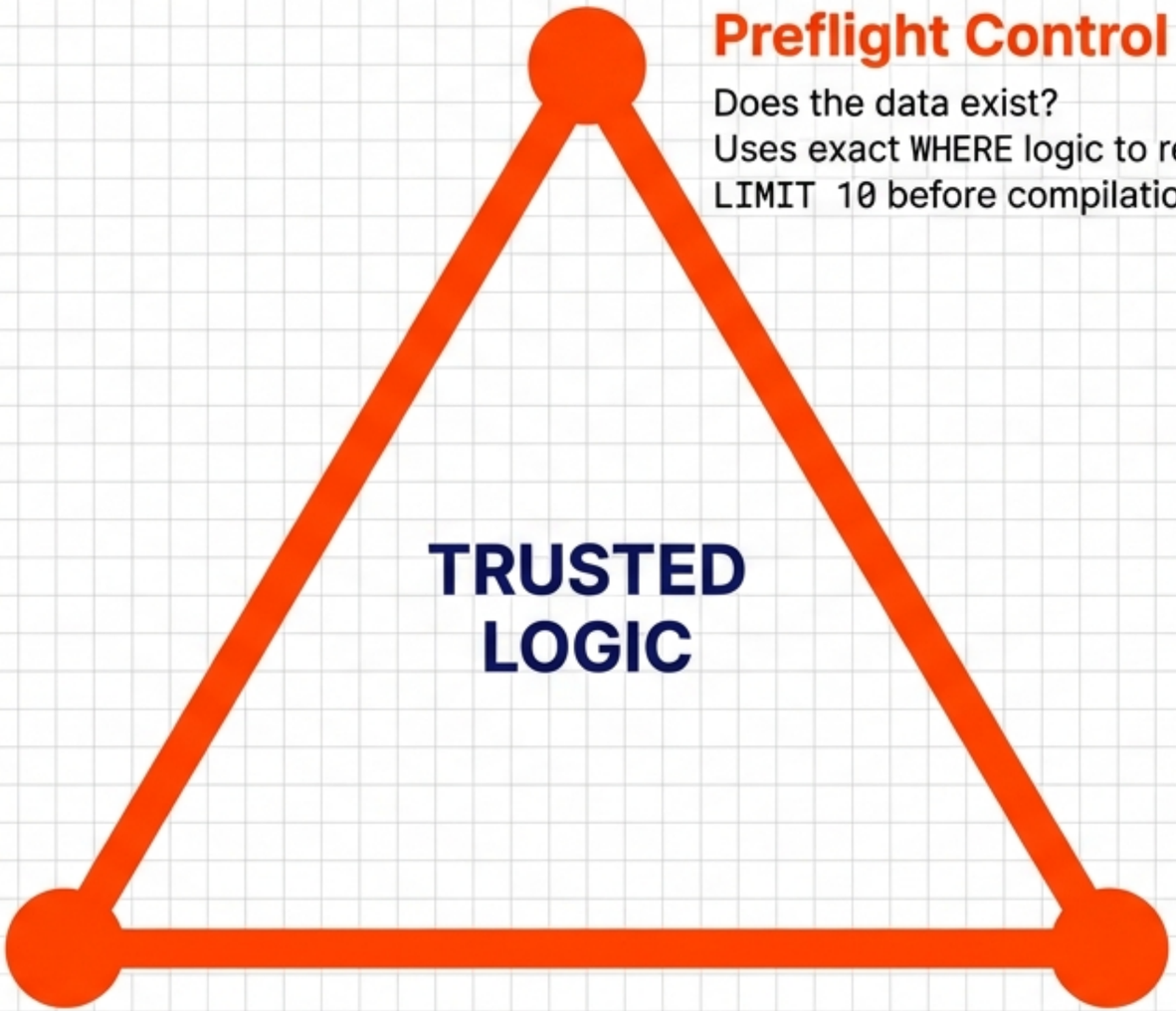
- 1-2 Minutes / Rule
- 100% Syntax Accuracy
- Accessible to SPARQL Users

The Validation Strategy

Ensuring data integrity through the Testing Triad.

Preflight Control

Does the data exist?
Uses exact `WHERE` logic to return `LIMIT 10` before compilation.



Negative Test

`WITHOUT DEFINE`.
Proves data is inferred at runtime and not materialized.

Positive Test

`WITH DEFINE input:macro-lib`.
Does the inference trigger correctly for known entities?

Business Applications of Dynamic Inference

CRM & Sales

Identify at-risk accounts in real-time without waiting for nightly batch jobs.

Executive Dashboards

Display global health metrics based on `NOW()` logic for up-to-the-minute reporting.

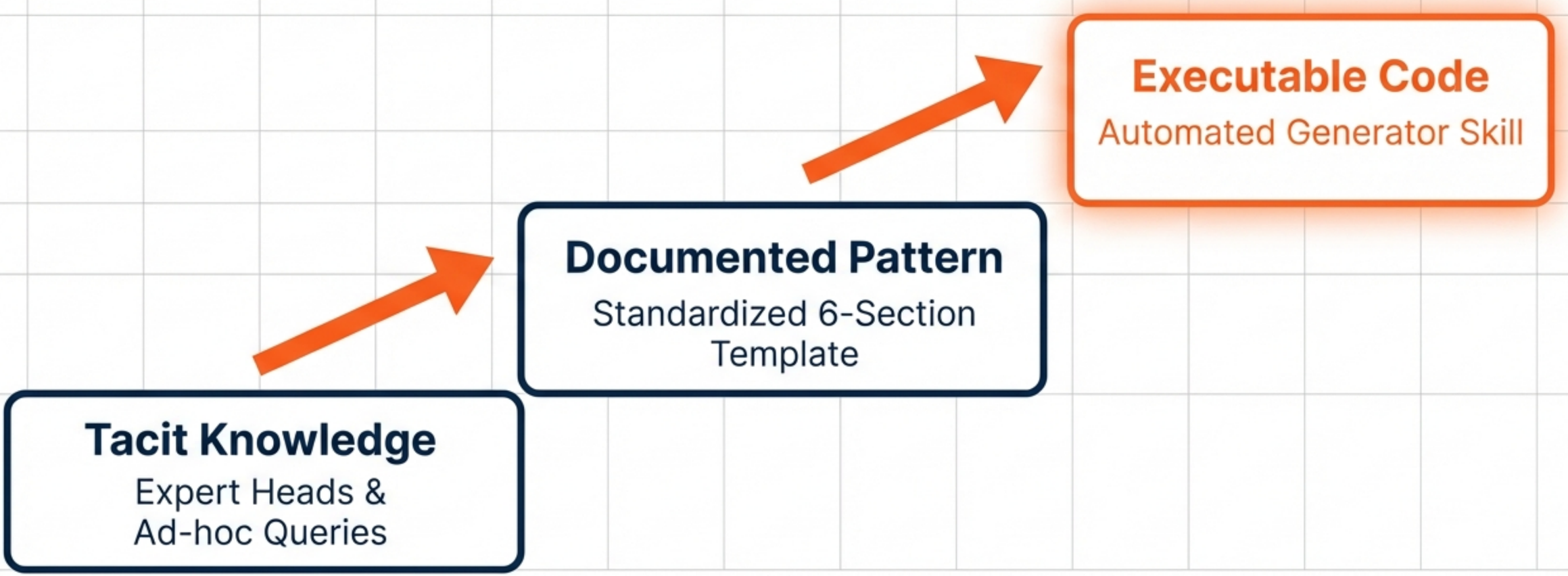
Marketing Precision

Target intersections like 'Dormant' + 'High Value' for win-back campaigns.

Trend Analysis

Run the same inference query against historical dates to track segment migration.

Strategic Value: From Tacit to Explicit



Result: Democratization. Non-database experts can now generate complex semantic rules using natural language.

Future Roadmap & Next Steps

Phase 1

Immediate (Now)

- Deploy Customer Analytics module
- Install Generator Skill (`spin-rule-gen`)

Phase 2

Expansion (Medium Term)

- Product & Geo-segmentation rules
- Enhance Skill with “Validation Mode”

Phase 3

Long Term Strategy

- Inference Governance (Version Control)
- Machine Learning Overlay

Sustainable, Scalable Inference

Inference is no longer a bottleneck—it is a scalable asset.

- The Module: 5-segment Customer Analytics system.

- The Tool: Automated Generator Skill (11KB).

- The Knowledge: Comprehensive documentation and pattern codification.

End of Presentation.