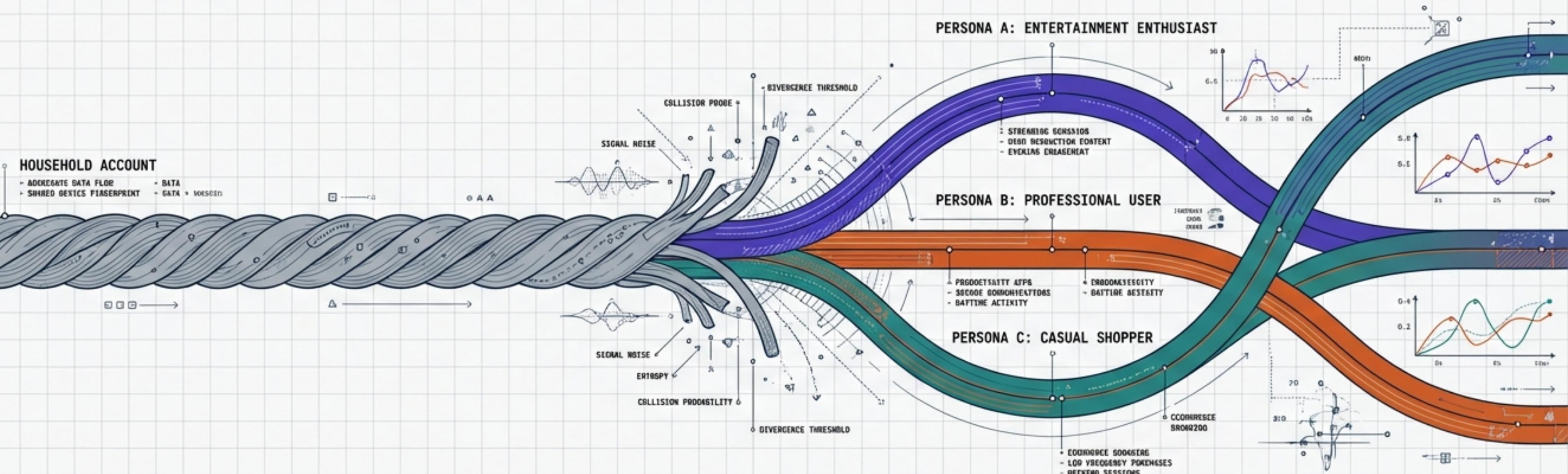


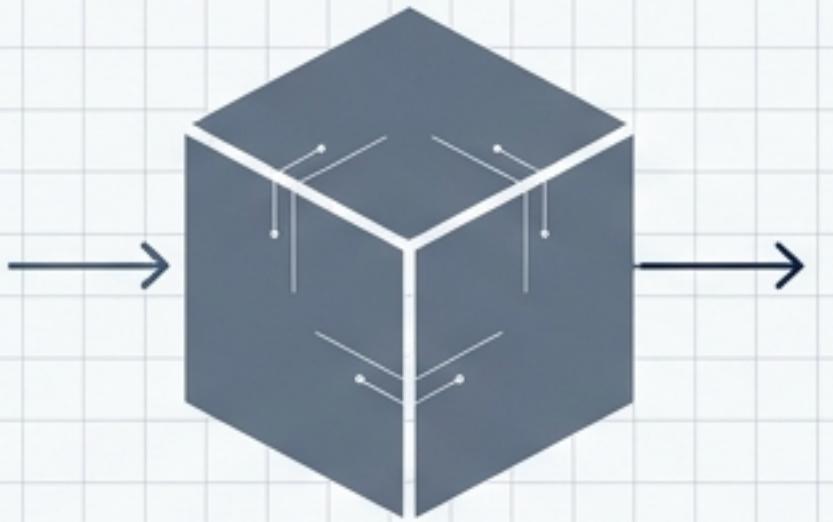
Multi-Platform Identity Resolution

A Probabilistic Framework for Solving the Shared-Account Attribution Challenge



Executive Summary

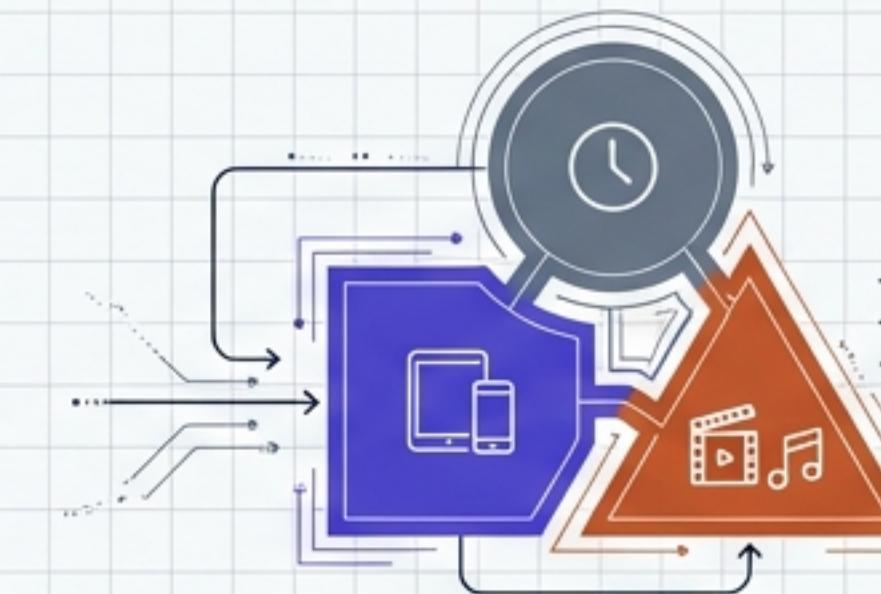
Helvetica Now Display



JetBrains Mono

~~The Challenge:~~ Streaming platforms view households as monolithic units. 40-60% of conversions are misattributed because One Subscriber ≠ One Viewer

Helvetica Now Display



JetBrains Mono

The Solution: A probabilistic engine utilizing behavioral fingerprinting (Time, Device, Genre) to infer distinct viewers without PII.



78%
Assignment Accuracy



22%
Attribution Lift



12M
Events/Hour Scale

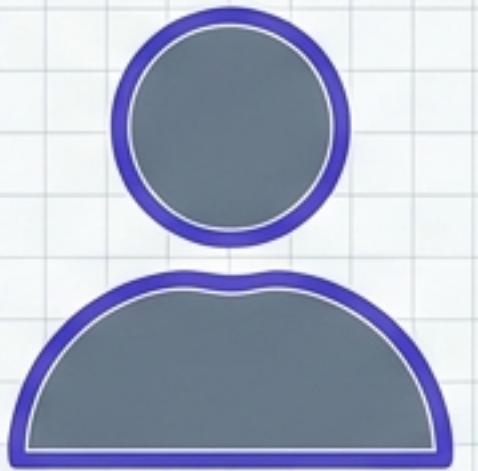


The Situation: The ‘Netflix’ Challenge

Reality vs. Metrics

The Metric

User Silhouette



1 Subscriber Account

The Reality

Primary Adult
(Drama)



Teen
(Action)



Child
(Cartoons)



← Attribution Failure.

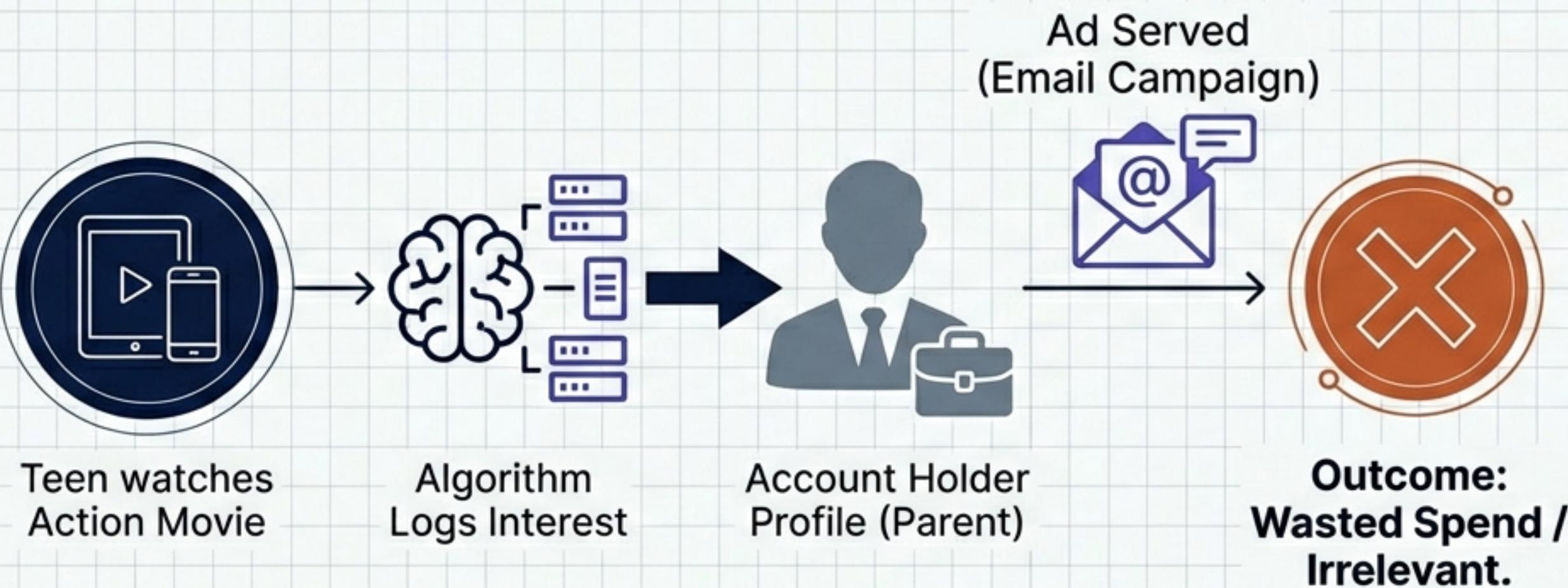
Email
Conversion

Social
Conversion

No
Conversion

Traditional analytics attribute all engagement to the bill-payer. We must shift from measuring Subscribers to measuring Monthly Active Viewers (MAV).

The Business Impact: Why Attribution Fails

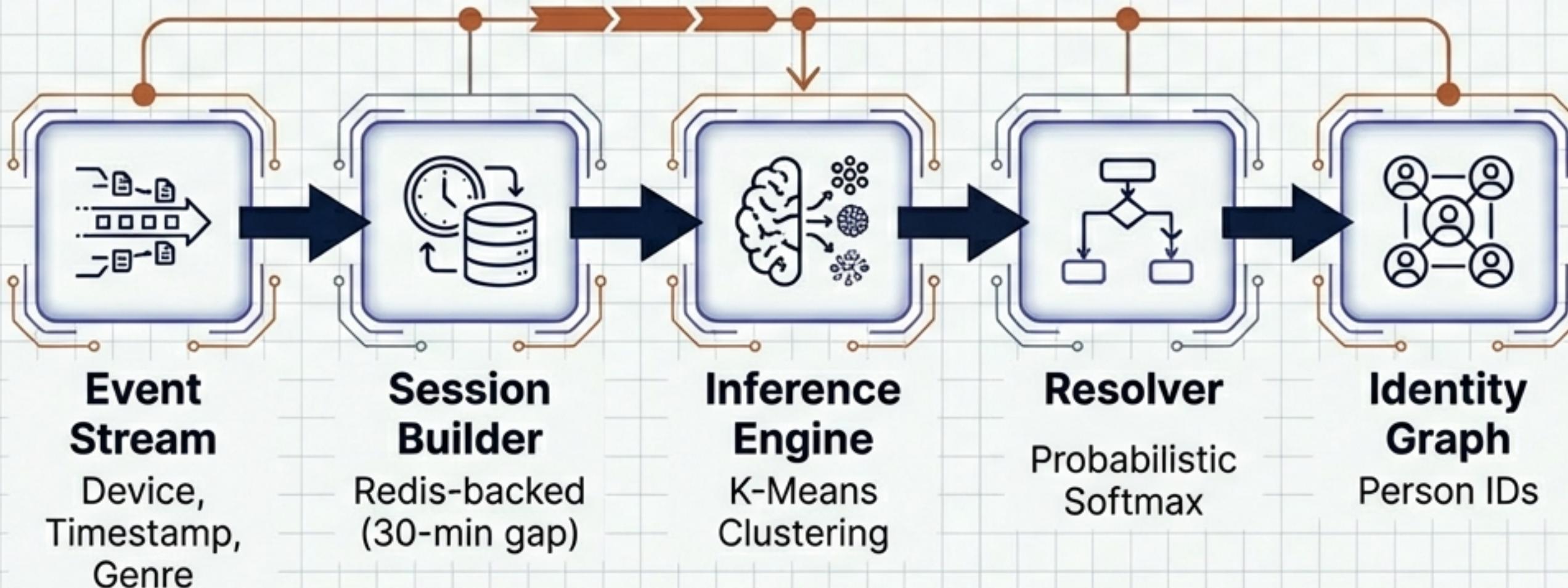


40-60%
Misattribution Rate.

Gap: Account-level data cannot answer 'Who watched?' or 'Who converted?'

Objective: Move attribution logic from 'Who pays?' to 'Who watches?'

Technical Solution: System Architecture



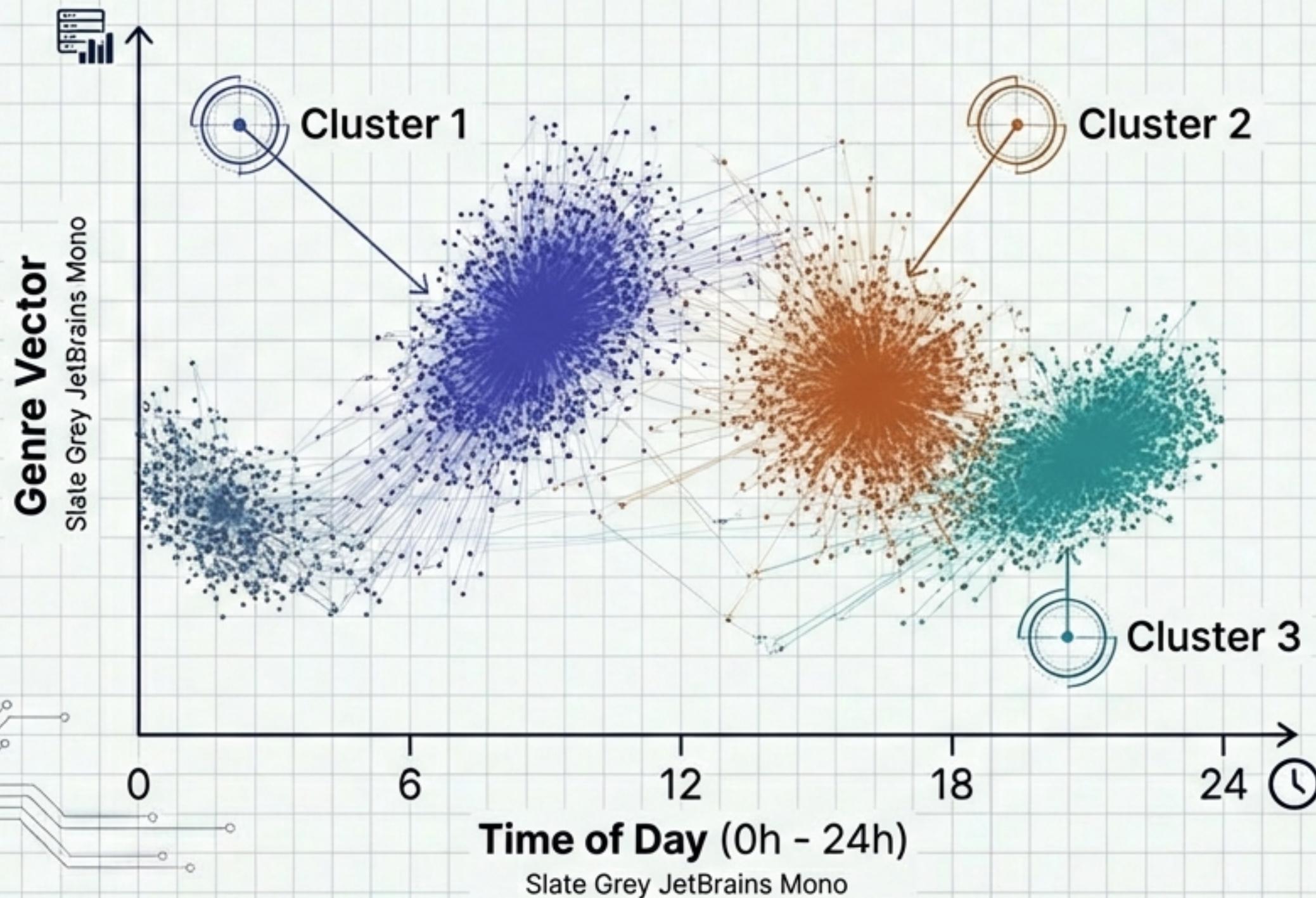
TECHNICAL SPECS

Latency: <100ms assignment

Throughput: 12M events/hour

Stack: Python, Redis, Incremental Clustering

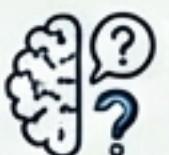
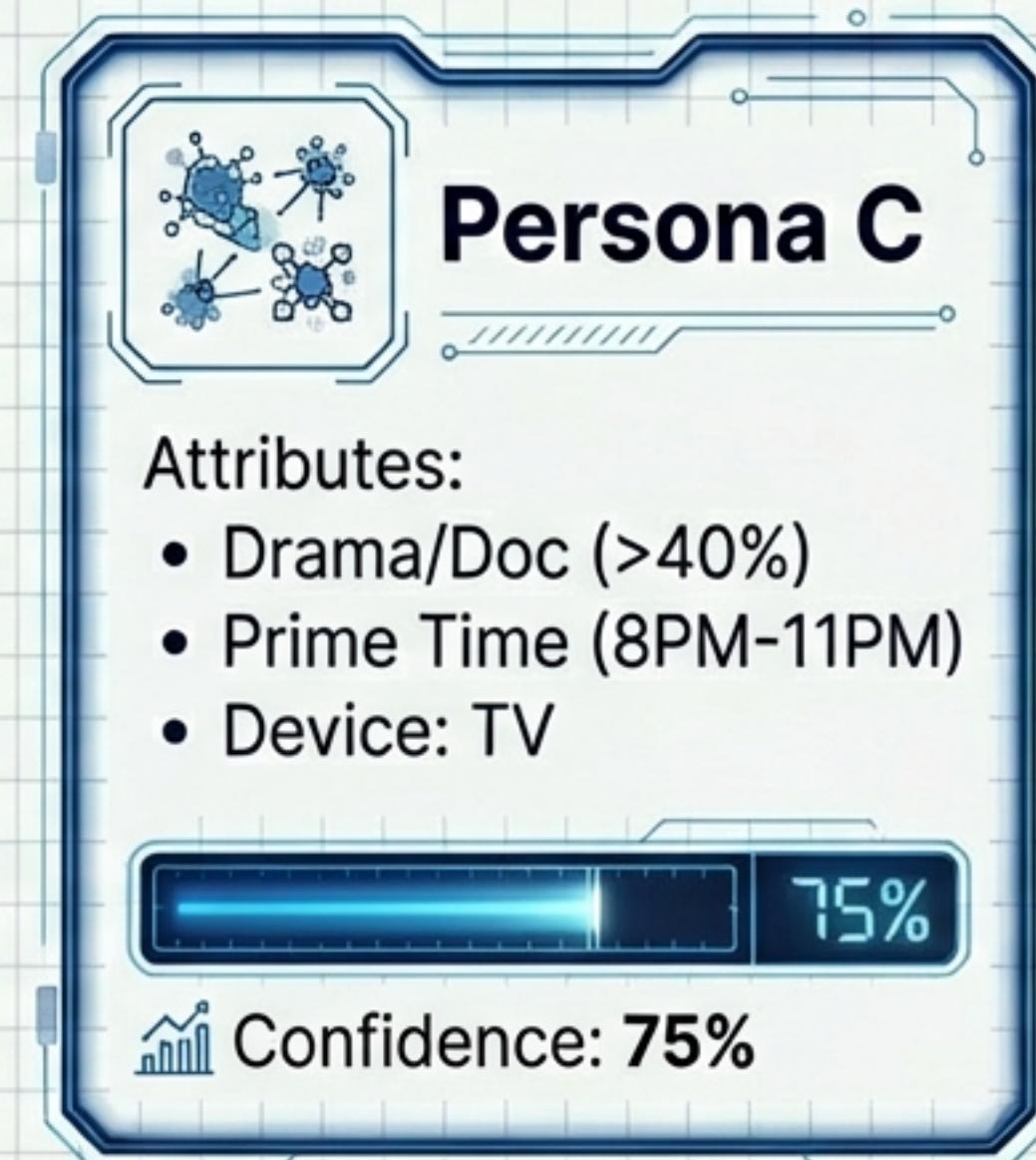
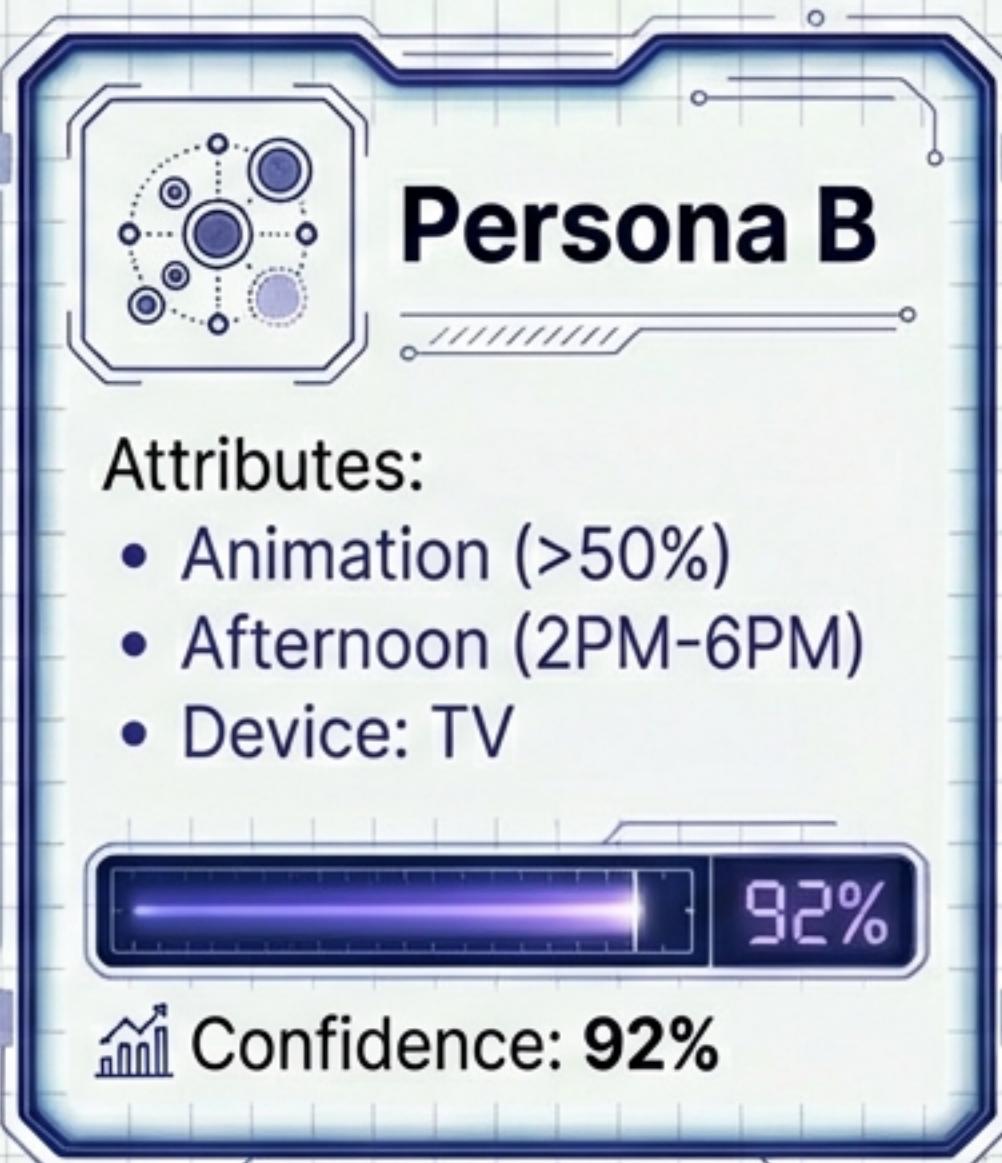
The Engine: Household Inference (Clustering)



TECHNICAL DETAILS

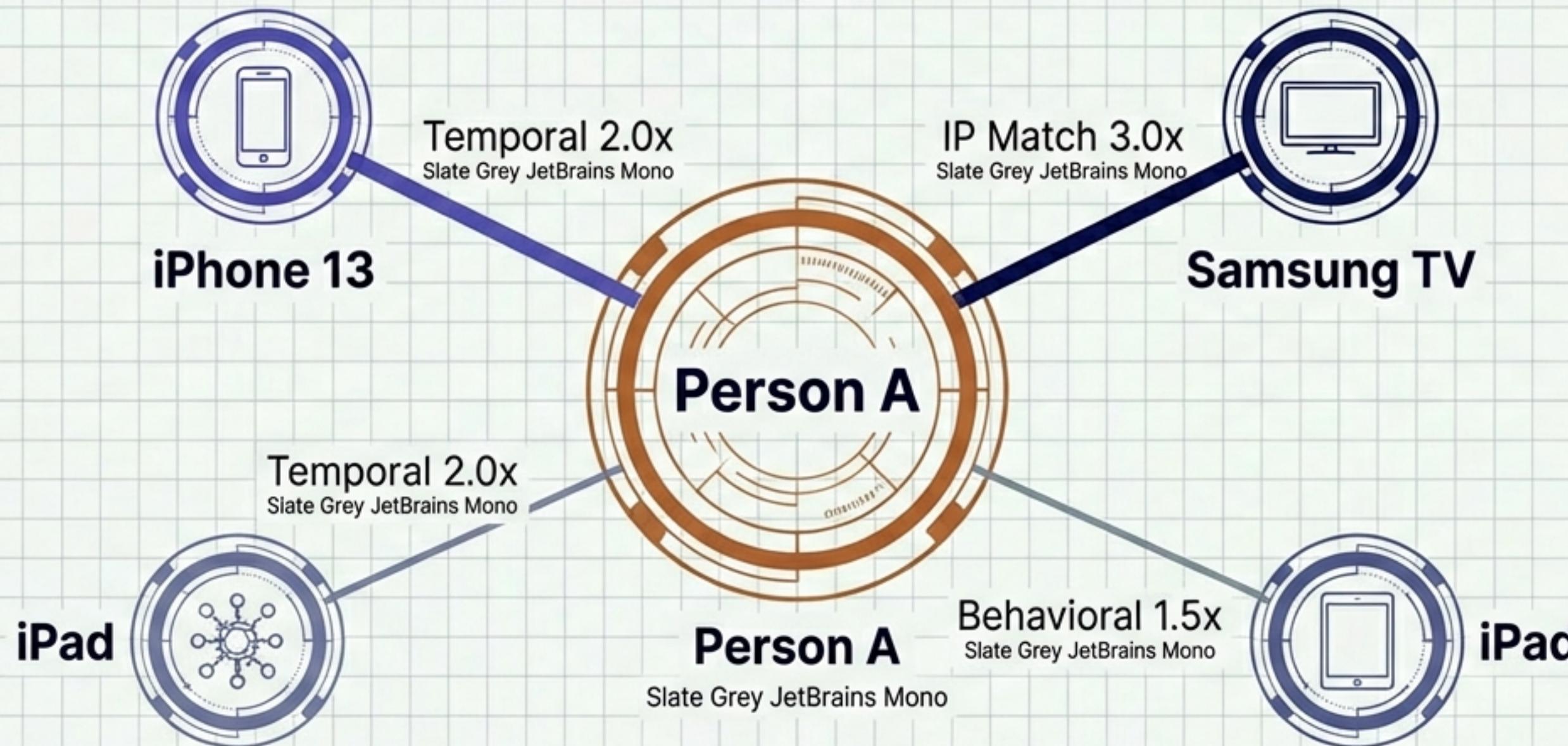
- Algorithm**
K-Means Clustering ($k=1$ to 6) with Silhouette Analysis.
- Logic**
Adaptive: If Silhouette Score < 0.3, default to single-user.
- Feature Vectors**
 1. Temporal (Hour/Day),
 2. Device (Type/Fingerprint),
 3. Content (Genre).

Personas & Profiles: Beyond the User ID



Key Insight: System acknowledges ambiguity. Output is not a hard label ('John'), but a **confidence score** ('80% Probability Person A').

Cross-Device Linking: The Unification Layer



Challenge

Slate Grey Helvetica Now Display

- Linking commute mobile sessions to home TV sessions.

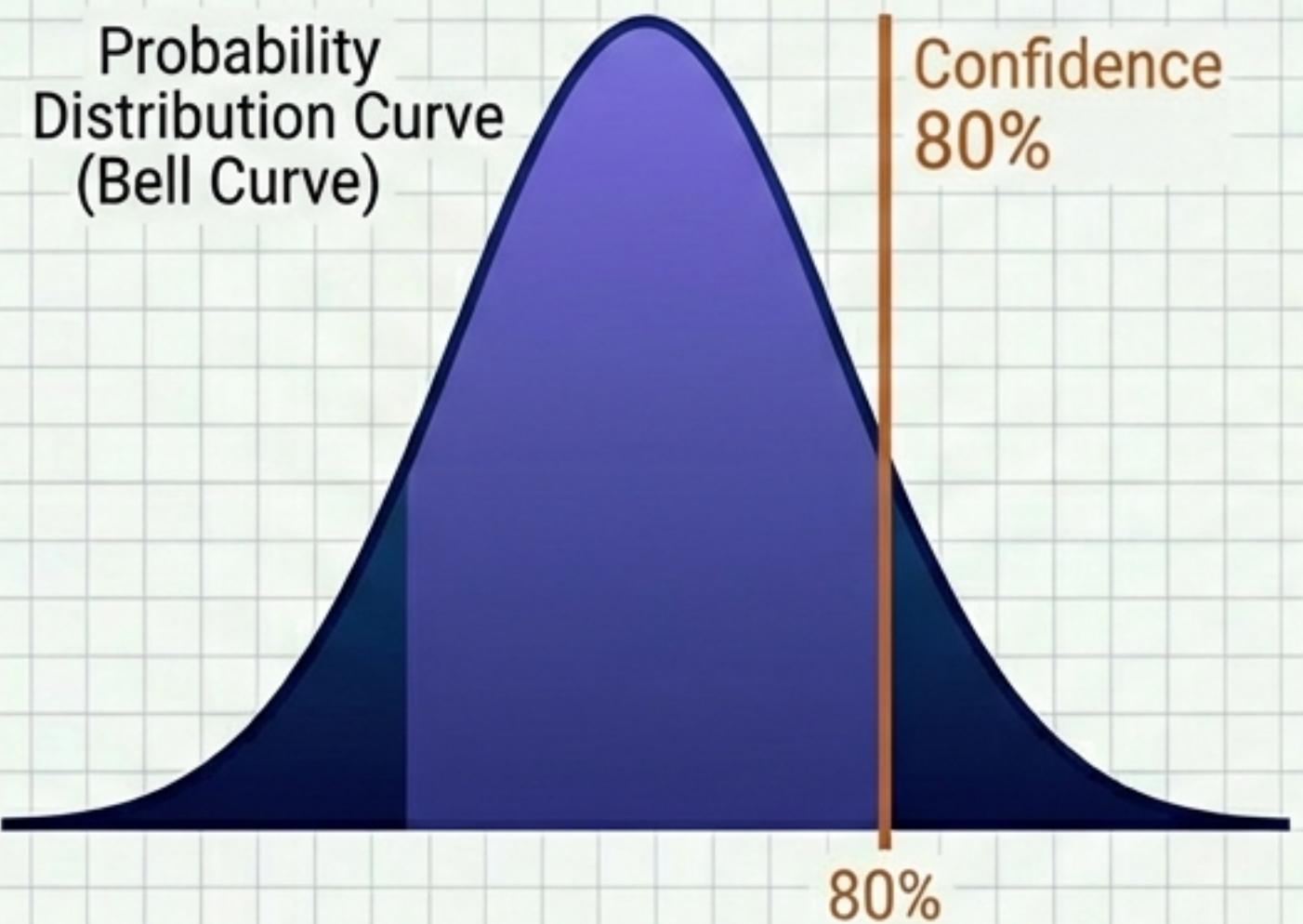
Logic

Slate Grey Helvetica Now Display

- Probabilistic Softmax constructs a device graph.
Connected components = Single Identity.

Probabilistic Assignment & Calibration

Probability Distribution Curve
(Bell Curve)



Confidence
80%

80%



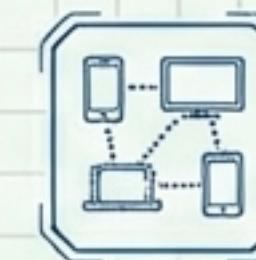
Soft Assignment

Sessions are assigned as a distribution (e.g., 0.8 Person A, 0.2 Person B).



Calibration

- Validated using Brier Score (<0.12).
- Ensures '80% confidence' is correct 80% of the time.



Co-Viewing Detection

- Triggers on:
 - TV Device
 - + Duration > 2 hrs
 - + Genre Shifts
- Distributes conversion value among co-viewers.

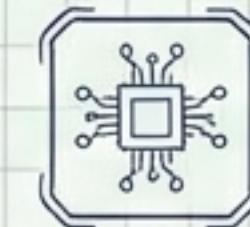
Probabilistic Models & Validation

Brier Score: <0.123
Slate Grey JetBrains Mono

Brier Score: <0.123
Slate Grey JetBrains Mono

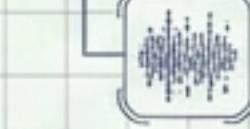
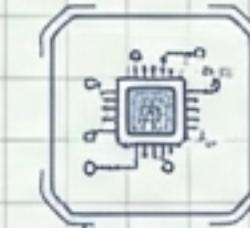
Data: Distribution vale per 0.1 Person
(, e.g., 0.8 Person A, 0.2 Person B)

Privacy & Security by Design



Constraint:

- Must function without names, emails, or explicit identifiers.



Methodology:

- **Data Hashing:** SHA-256 for all IDs.
- **Fingerprinting:** Behavioral patterns only.
- **Differential Privacy:** Noise injection ($\epsilon = 0.1$).



Compliance:

- **GDPR/CCPA Validated.** Right to deletion < 24hr.

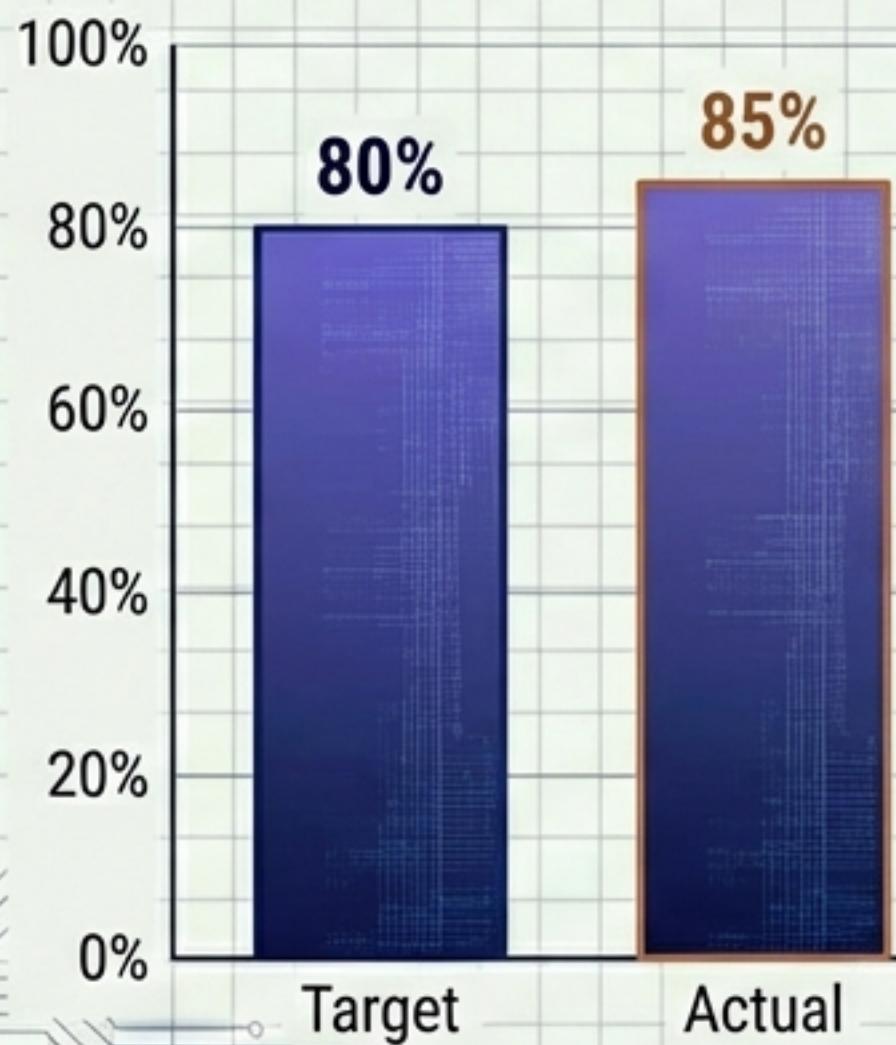
Validation: Synthetic Ground Truth

Tested against 50,000 synthetic profiles with known behaviors.

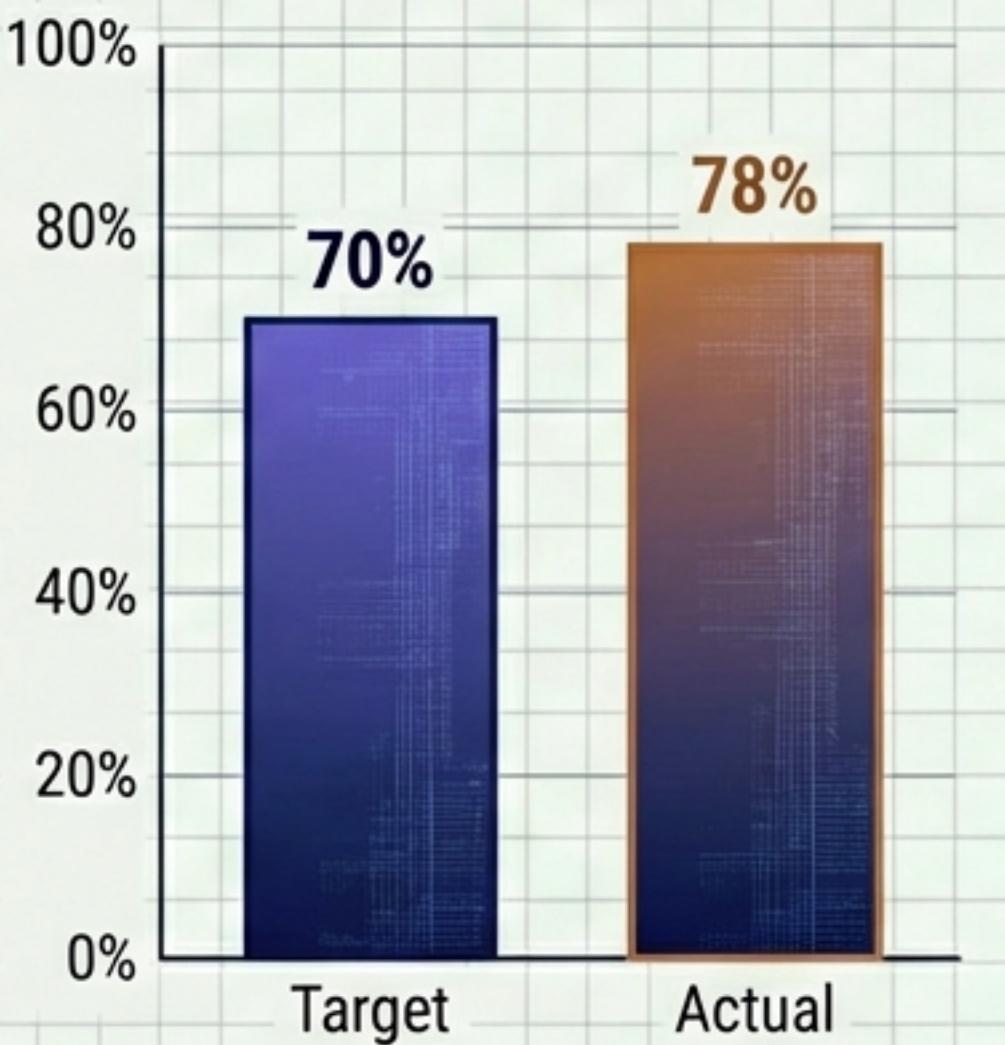
JetBrains Mono

Data Flow → Validated

Household Size



Person Assignment



Stress Test Panel

Device Sharing Resilience:
Handled 20% swapping.

Noise Resilience:
Resilient to 12% genre switching.

Cross-Device Linking
F1 Score: 82%.

Production Readiness: Benchmarks

System Dashboard

THROUGHPUT



12M events/hr



Status: Frozen v1.0.0

LATENCY P99

104ms

104ms

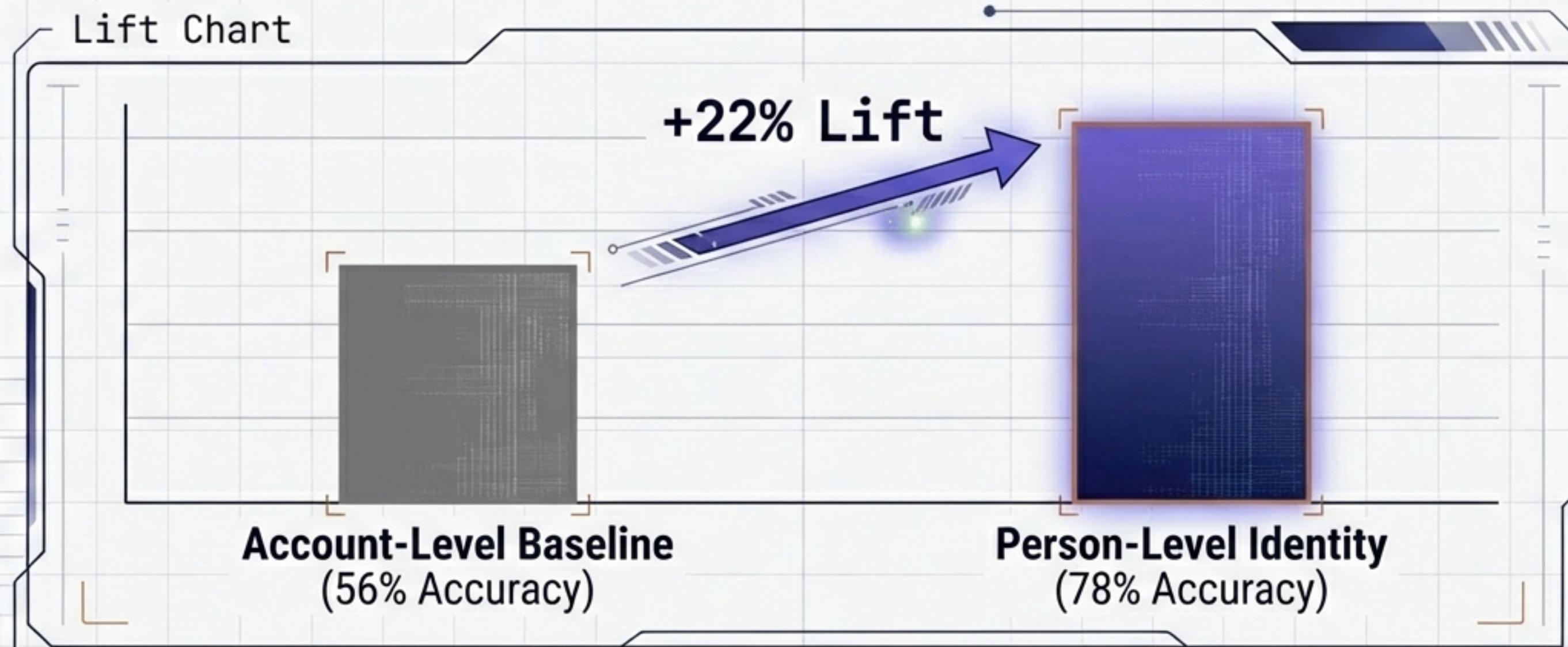
ERROR RATE

0.02%

↘ stable

running optimized python stack:
redis, incremental_clustering.py,
drift_detection.py

The 'So What?': Attribution Lift



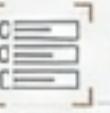
The Bottom Line: Knowing **who** converted changes the ROI calculation.

Strategic Value: Enables true personalization and accurate budget allocation.

Limitations & Future Roadmap

Timeline Analysis

CURRENT CONSTRAINTS

- **No Legal Identities**
(Cannot identify 'Mom').
- **Cold Start**
(Requires 10+ sessions). 
- **Correlation Only**
(No causal inference).   

FUTURE WORK

- **Phase 1:** Real-time online clustering.
- **Phase 2:** Real-time hierarchical clustering.
- **Phase 3:** Deep Learning personas. 

T=0

T=Future

Timeline Analysis // Version 1.0 to 3.0

Conclusion & Implementation

JetBrains Mono

```
> provision infrastructure --redis --kafka  
> deploy --image identity-resolution:v1.0.0  
> status check...
```

[SUCCESS] 78% Accuracy | 22% Lift | GDPR Compliant
[READY] System listening on port 8080...

JetBrains Mono

The Reference Implementation is frozen and ready.
Attribute to the **Person, not the Account.**