

1 Introduction

Aviation disruptions cost \$67.5 billion annually, yet no tool helps travelers predict delays before booking. Current flight tracking tools show only live status and lack the predictive foresight travelers need during the planning stage.



Travelers lack predictive foresight during planning

- Current tools show live status only, with no historical analysis or prediction for route planning
- Travelers cannot identify which routes, carriers, or time windows are delay-prone
- RouteRisk bridges this gap with interactive visualizations powered by 13.9M BTS flight records

2 Data

13.9M records 355 airports 5,969 routes Jan 2023 to Sep 2024

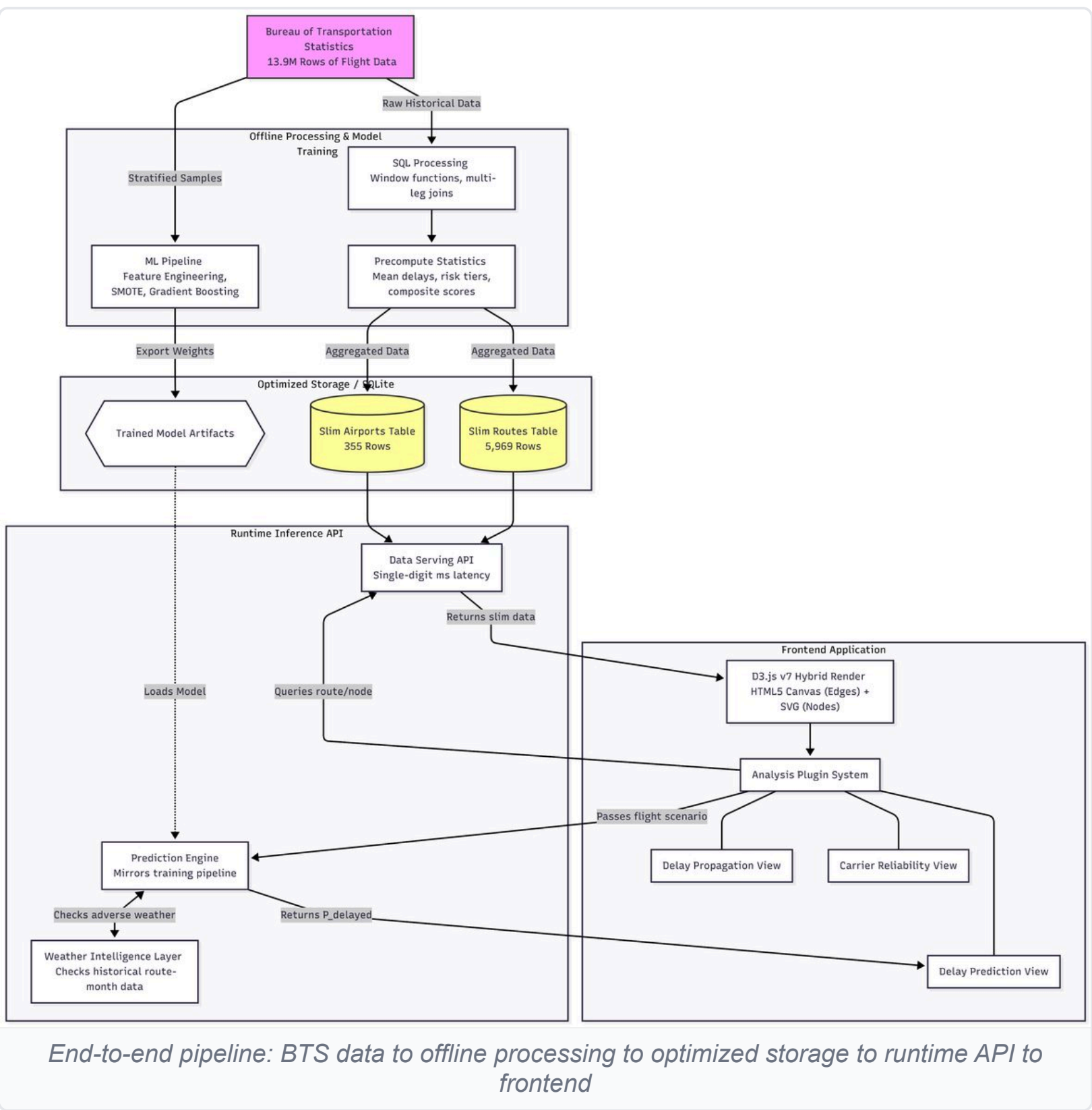
- Source: BTS Marketing Carrier On-Time Performance, direct download
- Features: tail_number chains, holiday proximity, binned departure periods
- Optimization: 13.9M rows refined into slim tables (355 airports, 5,969 routes) for single-digit ms latency

Processing Pipeline

BTS Raw → SQL → Features → Slim Tables → API

- Characteristics: Longitudinal time-series data spanning January 2023 through September 2024
- Feature engineering leverages tail_number to model inherited delay propagation; temporal features capture seasonal and time-of-day patterns

3 System Architecture

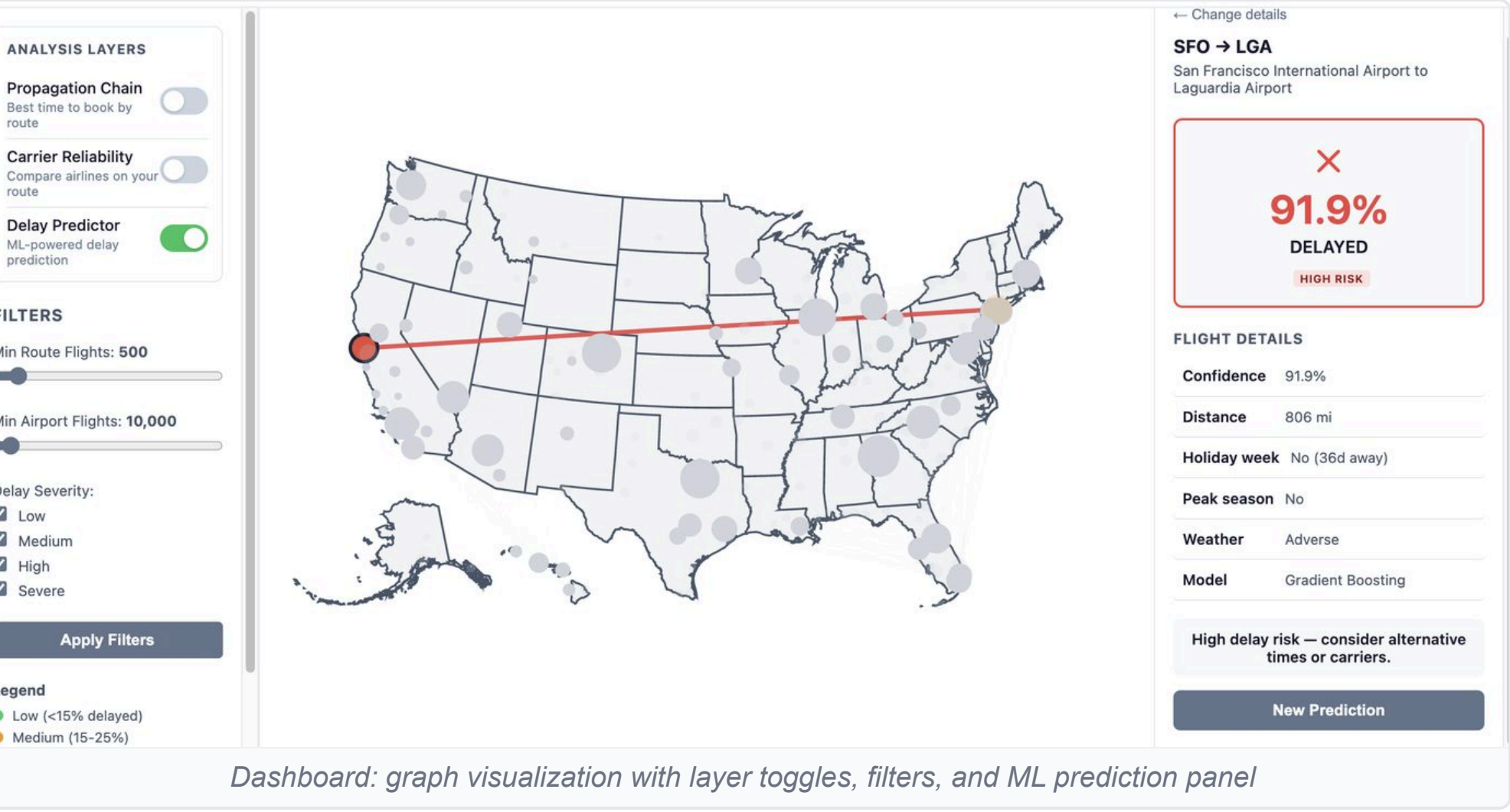


- Offline pipeline processes 13.9M BTS records through SQL aggregation and feature engineering into optimized slim tables
- Runtime API serves precomputed route statistics and ML predictions with single-digit millisecond latency
- Frontend layers (Propagation, Reliability, Prediction) share a unified D3.js graph control with independent toggle state

4 Approach

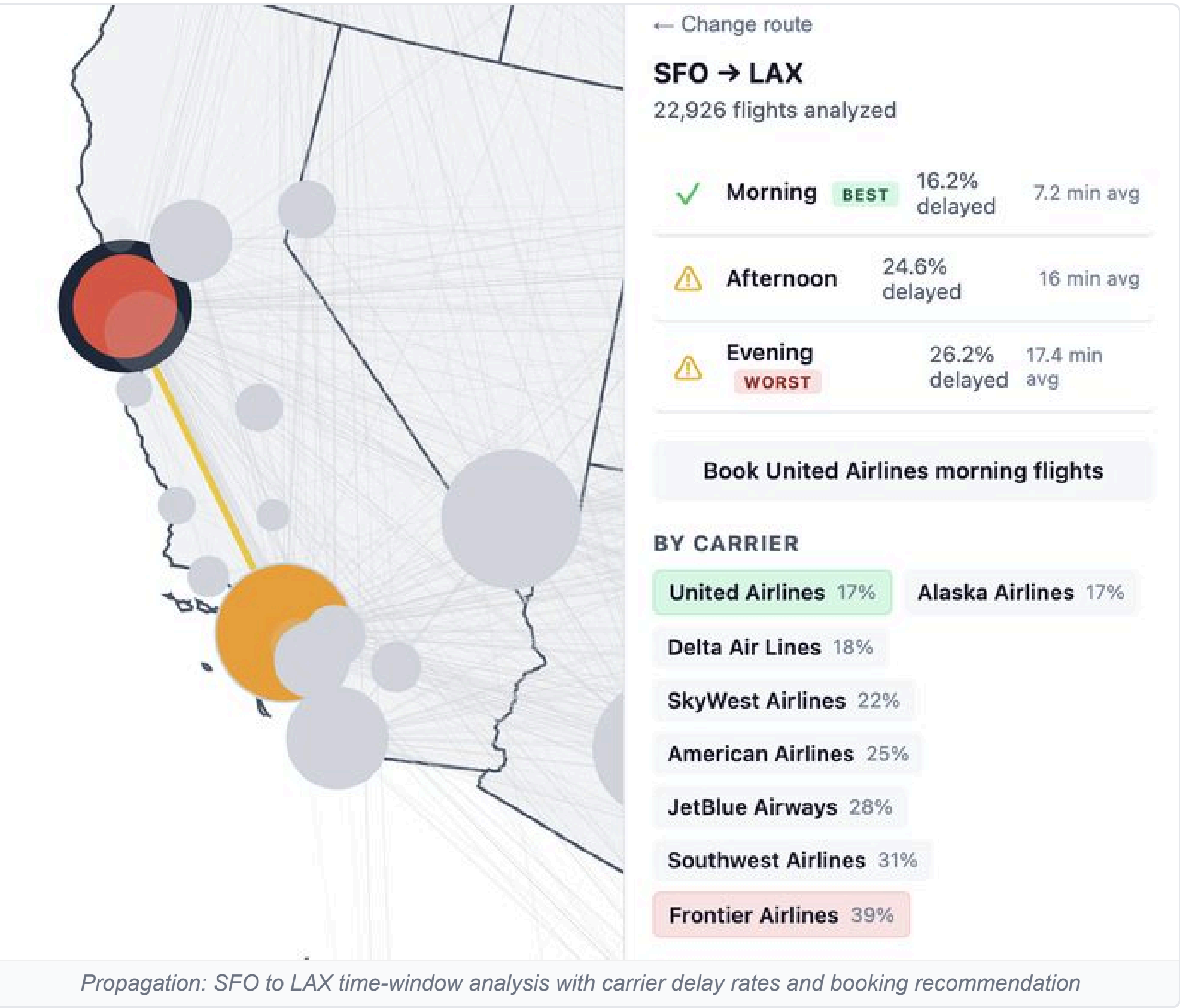
INNOVATION 1 Hybrid Canvas-SVG Rendering Engine

D3.js v7 renders 6,000+ edges on Canvas for performance; 355 nodes as interactive SVG circles. Edge opacity uses log scale from flight count; node radius uses sqrt scale. Positions computed once and cached.



INNOVATION 2 Delay Propagation via Tail Numbers

BTS tail_number traces inherited delays across multi-leg flights. SQL window functions order by (tail_number, date, time), bucket into morning/afternoon/evening. Route-window pairs (30+ observations) classified for upstream/downstream risk with booking verdicts.

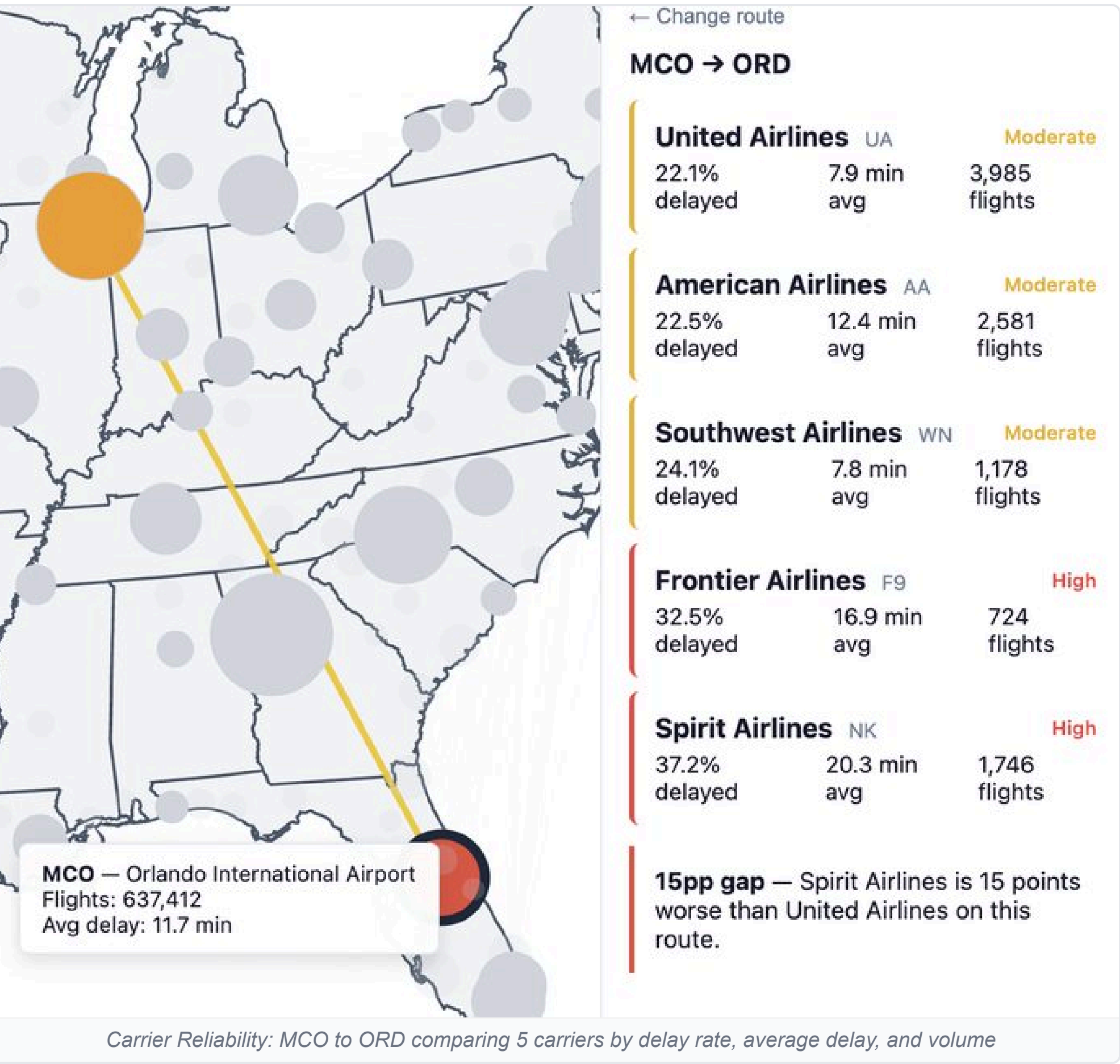


INNOVATION 3 Composite ML Pipeline

Binary classification (over 15 min late). 500K stratified samples, SMOTE, Gradient Boosting (200 trees, learning rate 0.05, depth 4). Selected via F1/ROC-AUC across 5 classifiers. API replicates training-time features.

INNOVATION 4 Plugin-Based Layer Architecture

Propagation, Carrier Reliability, and Prediction layers share the graph as a unified control. Users toggle layers without context-switching.



5 Evaluation & Results



Model Comparison (100K test)

Model	F1	ROC-AUC
Majority Baseline	0.000	0.500
Logistic Regression	0.778	0.927
Random Forest	0.812	0.930
Gradient Boosting ✓	0.813	0.931
XGBoost	0.774	0.932

Temporal Holdout (Oct to Dec 2024)

Class	Prec.	Recall	F1
On-time	0.94	0.98	0.96
Delayed	0.89	0.70	0.78
Weighted Avg	0.94	0.94	0.93

Strategy

- Predictive: Temporal holdout, trained Jan 2023 through Sep 2024, tested on Q4 2024
- Performance: Canvas-SVG vs standard SVG yields 60 FPS at full scale
- Usability: 5 participants, SUS = 82 (threshold: 70)

Feature Significance

- Statistical analysis confirmed tail_number (propagation) and carrier_id as the two strongest predictors of flight delays
- Temporal features (holiday proximity, binned departure periods) provide additional predictive signal for seasonal patterns

TOP PREDICTORS
tail_number and carrier_id validated our feature engineering approach.

6 How We Compare

RouteRisk vs. commercial tools and prior academic work:



- FlightAware etc. are live-only, with no historical or predictive capability
- Academic work targets industry analysts, not general travelers
- RouteRisk uniquely combines visualization, prediction, and propagation analysis
- Google Flights and Expedia provide fare tracking but no delay probability or propagation data for route planning
- OAG and Cirium serve enterprise analysts with batch reports, not interactive exploration for consumers

7 Conclusion

- Historical flight data can power both interpretable visualizations and accurate delay predictions
- Gradient Boosting achieved 0.93 ROC-AUC on unseen Q4 2024; SUS score of 82 confirms usability
- Canvas-SVG engine maintains 60 FPS across the full 13.9M-record network
- Future work: real-time weather forecasts, international routes, user-reported causes

Key Contributions

- First tool to unify delay propagation tracking, carrier reliability analysis, and ML prediction in one interface for general travelers
- Novel tail-number feature engineering traces inherited delays across multi-leg flights, validated as the strongest predictor of disruptions
- Custom hybrid Canvas-SVG rendering architecture enables 60 FPS interaction on million-record networks

KEY INSIGHT
Tail-number propagation and carrier identity were the strongest predictors. Inherited delays are a primary disruption driver.