Tom Sawyer Software Announces New Data Streams 1.0 Beta

ETL Multiple Enterprise Data Sources Into a Governed Knowledge Graph for AI and Decision Intelligence

BERKELEY, CA, UNITED STATES, November 26, 2025 / EINPresswire.com/ -- Tom Sawyer Software, the leader in graph and data visualization technology, today announced a beta release of their new product offering, Tom Sawyer Data Streams, a schema-driven platform to extract, transform, and load (ETL) structured and unstructured data into a single, governed, query-ready

Tom Sawyer Data Streams reduces integration effort and delivers a complete, accurate picture for lineage, impact analysis, and operational decisions.

knowledge graph. Data Architects and AI Platform Engineers can use Data Streams to subscribe to Apache Kafka (or Confluent) topics sourced from databases, files, and APIs; apply transformations and filters; then run flows continuously to normalize, enrich, and link changing streams in real time. The resulting knowledge graph is persisted in a graph database for scalable sharing and downstream analysis.
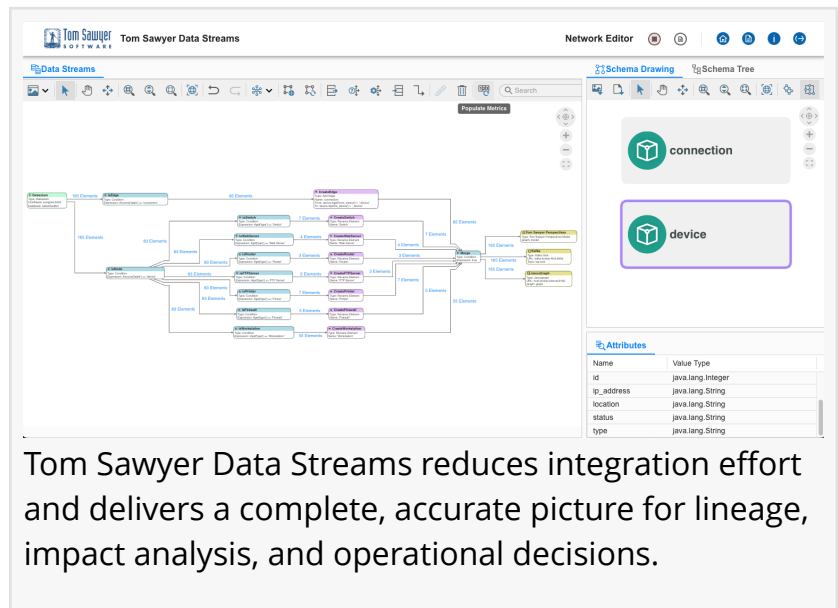
> Data Streams is a breakthrough capability for enterprises struggling with isolated and legacy datasets, costly migrations, and streamlining AI pipelines."
> *Brendan Madden, CEO of Tom Sawyer Software*

Data Streams works with existing data pipelines and catalogs in support of data preparation and processing within AI pipelines, creation of large language models, validating generative AI results, and integrating agentic results into a single-source-of-truth database. The platform lowers integration overhead, offering a unified view to support lineage, impact analysis, and operational decision-making.

"Data Streams is a breakthrough capability for enterprises struggling with isolated and legacy datasets, costly migrations, and streamlining AI pipelines," said Brendan Madden, CEO of Tom Sawyer Software. "Data Streams uses Kafka, CDC, and well-defined transformations to assemble a governed knowledge graph alongside your current stack. The result is lower storage and

migration spend, and a reliable context layer for operations and AI."

New in This Release:

Unified Streaming Data Integration: Data Streams subscribes to Apache Kafka (or Confluent) topics that users provision, then transforms and persists those events into a knowledge graph model, unifying legacy and streaming data in real time. Use standard tools, such as Kafka Connect, Change Data Capture (CDC) tools, or custom producers, to publish from the systems of record and subscribe in Data Streams. Users control which topics represent nodes, edges, or attributes, and secure access with existing Kafka authentication and encryption. Data Streams can consolidate disparate data from nearly any system—relational databases, graph databases, data warehouses, files, APIs, and more—into a single, governed, query-ready knowledge graph without expensive migrations and re-platforming.

Powerful Schema and Data Transformations: Data Streams automatically extracts schemas from Kafka topics, and users refine them in a visual editor including renaming fields, converting node types to edges, and applying advanced filters and rules with Spring Expression Language (SpEL). This provides users with control over normalization, enrichment, and linking so the resulting knowledge graph mirrors business semantics and remains consistent across sources.

Visual Data Flow Design: A web-based designer lets users define sources, transformations, conditions, and sinks to build and monitor data flows—accelerating delivery while reducing custom code. Automatic graph layout organizes complex pipelines for instant clarity, so users can iterate faster, validate logic visually, and promote changes with confidence, shortening the path from prototype to production.

Real-Time, Continuous Processing: Data flows can be run in batch or continuously to ingest and transform events as they arrive, keeping the knowledge graph always current. Low-latency processing supports operational decisions and analytics alike, ensuring downstream tools see the latest relationships and context.

Flexible Output and Storage: Transformed data is persisted into a graph database, and optionally users can publish results back to Kafka for downstream services. The resulting knowledge graph plugs directly into existing analytics, AI, visualization, and data science stacks—supporting retrieval-augmented generation and validation of generative outputs, as well as aggregating agentic results into a governed source of truth. Teams can query, explore, and operationalize insights without disruptive changes or costly data migrations.

Enterprise-Grade Security and Authentication: Data Streams supports authentication with OAuth 2.0 and Keycloak for secure, multi-user environments. It aligns with enterprise standards for access and encryption while maintaining a clean, auditable boundary around sensitive data.

Docker-Based, Air-Gapped Installation: A dedicated Docker installer streamlines deployment in

cloud, on-premises, and fully offline (air-gapped) environments. Operations teams can standardize installs and upgrades across clusters, meeting security constraints without sacrificing speed.

Seamless Integration with Tom Sawyer Graph and Data Visualization Tools: Data Streams pairs with Tom Sawyer Perspectives and Tom Sawyer Explorations to visualize the resulting knowledge graph from the data streams, providing automatic graph layouts, interactive filtering, and intuitive data exploration. Quickly moving from ingestion to interactive insight enables stakeholders to explore patterns, validate models, and share findings.

To learn more about Data Streams and its capabilities visit the product web page or request a live demo or free trial today.

About Tom Sawyer Software

Tom Sawyer Software is the leading provider of software and services that enable organizations to build highly scalable and flexible graph and data visualization and analysis applications. These applications are used to discover hidden patterns, complex relationships, and key trends in large and diverse datasets. Tom Sawyer Software serves clients with needs in link analysis; network topology; architectures and models; schematics and maps; and dependencies, flows, and processes. We help clients federate and integrate their data from multiple sources and build the graph and data visualization applications that are critical to analyzing and gaining insight into their data.

Caroline Scharf
Tom Sawyer Software
+15012084370 ext.
email us here
Visit us on social media:
LinkedIn
Instagram
Facebook
YouTube
X